An Approach to the Prediction of Chemosensory Effects of Volatile Organic Compounds on Humans

A Thesis presented to the University of London in partial fulfilment of the requirements for the degree of Doctor of Philosophy in the Faculty of Science

by

Ricardo Sánchez Moreno

Sir Christopher Ingold Laboratory
Chemistry Department
University College London

September 2007

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

To my mother

Abstract

The aim of this project was to construct equations to predict the mechanism of threshold values of sensory irritation (nose and eye irritation) caused by Volatile Organic Compounds (VOCs). Mathematical models were established to predict Nasal pungency thresholds (NPTs) and odour detection thresholds (ODTs).

Firstly, the Abraham Solvation Equation was used to determine the octanol/water partition coefficient (log P_{oct}) for fluorescein. A similar equation was used to characterize several Gas-Liquid Chromatography (GLC) stationary phases.

Abraham descriptors were determined for several VOC compounds using data obtained from experiments using GLC and data on water-solvent partition coefficients.

The QSAR analysis on NPTs and ODTs revealed that "selective" processes (i.e., chemically-broad, transfer driven effects) overwhelmingly dominate chemesthetic detection, whereas both selective and "specific" (i.e., chemically-narrow, ligand-receptor interactions) processes control olfactory potency. To understand further the nature of the chemical factors that influence ODT values, the possibility was explored that NPT values could be used to estimate "selective" effects in ODTs.

In the present study, concentration-detection, i.e., detectability, functions for the human olfactory detection of some chemical vapours of diverse chemical structures were analyzed. The functions cover the complete peri-threshold range from chance to almost perfect detection and provide much more information than can be conveyed by a single "threshold value". These odour functions have been modeled with a uniform approach: a simplified sigmoid equation, $y = 1 / (1 + e^{[-(x-C)/D]})$, and have been correlated with the set of up to five physicochemical descriptors taken from the Abraham solvation equation that has quite successfully described the simpler "odour threshold" and "nasal pungency" psychophysical outcome.

Previous studies of ocular and nasal chemesthetic thresholds along and across homologous chemical series have suggested the existence of a "cut-off", i.e., a point where a homolog fails to stimulate even at vapor saturation. All larger homologs fail as well. The present study sought, first, to identify the particular cut-off member along homologous alkylbenzenes and 2-ketones and, second, to probe the likely basis for the effect.

Table of Contents

Abstract i

Table of contents ii

List of tables ix

List of figures xiv

Ethical approval xix

Abbreviations xx

Acknowledgements xxii

Chapter 1 An introduction to Indoor Air Pollution

1.0. Introduction 1

1.1. Sources of Indoor Air Pollutants 2

1.1.1. Volatile Organic Compounds 3

1.2. Assessment of Exposure 5

1.2.1. Personal Monitoring 6

1.2.2. Biological Monitoring 6

1.2.3. Controlled Exposure Chamber 7

1.3. Concentration of Indoor Air Pollutants 7

1.4. Indoor Air Pollutants and Health 10

1.4.1. Multiple Chemical Sensitivity 11

1.4.2. Building-Related Illness 11

1.4.3. Sick Building Syndrome 11

1.4.4. Volatile Organic Compounds and Health Effects 13

1.5. References 15

Chapter 2 Study of Odour and Nasal Pungency

2.0. Introduction 19

2.1. Sensory Channels 21

2.1.1. Olfaction 21

2.1.2. Chemical Senses 34

2.2. Overview of Odour and Nasal Pungency Measurement Techniques 38
2.2.1. Static System 40
2.2.2. Dynamic System 41
2.2.3. Environmental Chambers 41
2.2.4. Methodology 43
2.3. Psycophysical Responses to Chemical Mixture 46
2.4. References 47

Chapter 3 Study of Eye Irritation Thresholds

3.0. Introduction 54
3.1. Sources and Pathway for Eye Irritation 55
3.2. Anatomy and Physiology of The Eye 57
3.3. Eye Irritation Measurements 60
3.3.1. Eye Irritation Detection Threshold 61
3.3.2. Eye Irritation Psychometric Functions 62
3.4. Psycophysical Responses to Chemical Mixture 62
3.5. References 63

Chapter 4 An introduction to the Abraham Solvation Equation

4.0. Introduction 67
4.1. Linear Free Energy Relationship 68
4.1.1. Solvation Parameter Model 68
4.1.2. The Solvatochromic Comparison Approach 69
4.2. The Abraham Solvation Equation 72
4.2.1. The Abraham Solute Parameters 72
4.2.2. Applications of the Abraham Solvation Equation 86
4.3. References 87

Chapter 5 An introduction to Gas Liquid Chromatography

5.0. Introduction 92
5.1. Retention in Gas-Liquid Chromatography 94
5.1.1. Influence of Mobile Phase Physical Properties 95
5.1.2. Property Estimations 97

5.2. Retention Index System and Their Applications 98
5.2.1. Retention Index Standards 99
5.2.2. Temperature Dependence of Kovats Indices 99
5.2.3. Unified Retention Index 101
5.2.4. Connection between the Physicochemical Quantities and Retention Indices 102
5.3. Qualitative and Quantitative Analysis 103
5.3.1. Qualitative Analysis 103
5.3.2. Quantitative Analysis 104
5.4. Characterisation of Stationary Phases 105
5.4.1. Introduction 105
5.4.2. Rohrschneider's and McReynolds' Phase Constants 105
5.4.3. Solvation Parameter Model 107
5.5. References 111

Chapter 6 An introduction to Partition Coefficients

6.0 Introduction 114
6.1 Theoretical 115
6.1.1 Henry's Law 115
6.1.2 Nonideal Behavior of Solutes 115
6.1.3 Thermodinamics of Partitioning Systems 117
6.1.4 Energy Requirements for Phase Transfer 121
6.2 Relationships between Gas to Solvent and Gas to Water Partitions, Water to Solvent Partitions, and Solubilities 122
6.3 Experimental Methods 124
6.4 References 125

Chapter 7 A Brief Introduction to Statistical Analysis

7.0. Introduction 127
7.1. Multiple Linear Regression Analysis 127
7.1.1. Linear least squares 127
7.1.2. Interpreting the results 128
7.2. Paired T-Test 133

7.3. Principal Components Analysis 134
7.3.1. Eigenanalysis 135
7.3.2. Analysis of the eigenvalue and score plots 136
7.4. Training Set and Test Set 136
7.5. References 137

Chapter 8 Aims of the Present Work

139

Chapter 9 Determination of the Water/Octanol Partition Coefficient as the log P value for Fluorescein and Diacetylfluorescein

9.0. Aim 141

I Fluorescein

9.1. Introduction 141
9.2. Method 142
9.2.1. General analysis 142
9.2.2. Determination of the Partition Coefficient 144
9.2.3 Experimental 146
9.3 Results 148
9.4 Calculation of descriptors for Fluorescein 150
9.5 Conclusion 151

II Diacetylfluorescein

9.6 Method 154
9.6.1 General analysis 154
9.6.2 Experimental 155
9.7 Results 156
9.8 Conclusion 159
9.9 References 160

Chapter 10 Characterization of GLC stationary phases by Abraham's method

10.0. Introduction 161

10.1. Experimental 162
10.1.1. Materials 162
10.1.2. Column preparation 163
10.1.3. Instrumentation 164
10.1.4. Retention studies 164
10.2. Results and Conclusions 166
10.2.1. Multiple linear regression analysis 166
10.2.2. Results 168
10.2.3. Interpretation of LSER coefficient values 172
10.3. References 173
10.4. Appendix 174

Chapter 11 Determination of Solvation Properties of VOCs from GLC data

11.0. Introduction 179
11.1. Methodology. The System of Abraham 181
11.1.1. The E and V descriptors 181
11.1.2. The L, A and S descriptors 181
11.1.3. The B descriptor 183
11.1.4. A general method for the determination of descriptors 183
11.2. Results 185
11.2.1. Statistical analysis 190
11.3. References 194
11.4. Appendix 197

Chapter 12 A New Algorithm for Nasal Pungency and Odour Detection Threshold based on the Abraham Solvation Equation

12.0. Introduction 206
12.1. Methodology 207
12.2. Results 208
12.2.1. Algorithm for Nasal Pungency Thresholds 208
12.2.2. Algorithm for Odour Detection Thresholds 213
12.2.3. Initial analysis 219
12.3. Estimation of Nasal Pungency and Odour Detection Thresholds 220

12.4. Conclusion 231
12.5. References 232

Chapter 13 A Quantitative Structure-Activity Analysis on the Relative Sensitivity of the Olfactory and the Nasal Trigeminal Chemosensory Systems

13.0 Introduction 234
13.1 Materials and Methods 235
13.2 Results and Conclusions 237
13.3 Conclusions 243
13.4 References 244
13.5 Appendix 247

Chapter 14 An Algorithm for 323 Odour Detection Threshold Based on the Nagata Data Set

14.0. Introduction 250
14.1. Methodology 251
14.2. Results and Discussion 252
14.2.1. Correlation of the Nagata data set 252
14.2.2. Comparison of the Nagata and Cometto-Muñiz and co-workers data set 263
14.2.3. Incorporation of the Hellman and Small data set 272
14.3. References 277

Chapter 15 Concentration-detection Functions for the Olfactory Detectability of Airborne Chemicals

15.0. Introduction 280
15.1. Materials and Methods 281
15.1.1. Subjects and stimuli 281
15.1.2. Procedure 282
15.1.3. Data Analysis 283
15.2. Results 283
15.3. Discussion 287
15.4. References 296

Chapter 16 Concentration-detection Functions for Nasal Pungency of Airborne Chemicals

16.0. Introduction 299

16.1. Materials and Methods 300

16.1.1. Subjects 300

16.1.2. Stimuli and Equipment 300

16.1.3. Procedure 301

16.1.4. Data Analysis 301

16.2. Results 302

16.3. Discussion 304

16.4. References 307

Chapter 17 Chemical Boundaries for Detection of Eye Irritation in Humans from Homologous Vapours

17.1. Introduction 309

17.2. Experiment 1: Detection of Eye Irritation at Room Temperature (23 °C) 309

17.2.1. Materials and Methods 309

17.2.2. Results 311

17.3. Experiment 2: Comparison of the Eye Irritation Detectability of Heptyl benzene and 2-tridecanone at 23 and at 37 °C 313

17.3.1. Materials and Methods 313

17.3.2. Results 314

17.4. Discussion 316

17.5. References 326

Chapter 18 Suggestions for Future Work

329

List of tables

Table 1.1	Outdoor sources of major indoor air pollutants	3
Table 1.2	Sources of common volatile organic compounds in indoor air	4
Table 1.3	Summary of Indoor and Outdoor Air concentrations and Exposures	8
Table 1.4	Health Effects of selected Volatile Organic Compounds	13
Table 1.5	Effects of Human Health and Well-being Caused by Exposure to VOC in Indoor Climate	14
Table 2.1	Characteristics of the Term "Chemical Sensory Irritation"	35
Table 4.1	Old and new notation of the Abraham solute descriptors	73
Table 4.2	Atom contributions for calculation of **V** (in cm^3 mol^{-1})	75
Table 4.3	A comparison between α_2^H and **A** values	78
Table 4.4	Comparison between β_2^H and **B** values	79
Table 4.5	Water-solvent and gas-solvent processes used in the determination of solute descriptors	81
Table 4.6	Table of the available solute descriptors	82
Table 5.1	Comparison of manual methods for determining peak areas	104
Table 5.2	Characteristic solute interactions associated with the first five McReynolds' probes (Rohrschneider's probes in parentheses)	107
Table 5.3	Stationary phases and McReynolds values	108
Table 5.4	System constant derived from the solvation parameter model for packed column stationary phases (120-122°C)	110
Table 6.1	Temperature effect on log P	119
Table 6.2	A few values of Q at 298.15K	123
Table 9.1	Estimation of the values of the fluorescein descriptors	143
Table 9.2	Coefficients for water/solvent partitions at 298 K. Calculated values for fluorescein, using the estimated descriptors in Table 9.1	143
Table 9.3	Systems studied	146

Table 9.4	Results found in this work. Observed absorbance and partition coefficient	148
Table 9.5	Summary of the observed values in this work	149
Table 9.6	Calculated values of log P for fluoroscein by the Solver method; the corresponding obtained descriptors are in Table 9.7. The observed values are from Table 9.5	150
Table 9.7	Results found for the solute descriptors using the seven equations in Table 9.5 and the total set of 15 equations	151
Table 9.8	Comparison between the value found in this study and the values found in the literature	151
Table 9.9	Estimation of the values of the diacetylfluorescein descriptors	154
Table 9.10	Coefficients for water/solvent partitions at 298 K. Calculated log P (log P_{calc}) values for diacetylfluorescein in different systems, using the estimated descriptors in Table 9.9	155
Table 9.11	Solutions prepared to study the water/cyclohexane system	156
Table 9.12	Observed absorbance obtained for the water/cyclohexane system using the solution A	157
Table 9.13	Comparison of calculated values for diacetylfluorescein with the observed value by Eric Clarke	158
Table 10.1	Summary of log t_{rel} values for on squalane	166
Table 10.2	Summary of log t_{rel} values on Carbowax 20M	167
Table 10.3	Summary of log t_{rel} values on DEGS	168
Table 10.4	Linear Free Energy Relationship (LFER) coefficients	168
Table 10.5	Comparison of LFER coefficients for squalane	169
Table 10.6	Regression Statistics for Eq. 10.4 for squalane	170
Table 10.7	Comparison of LFER coefficients for Carbowax 20M	170
Table 10.8	Regression Statistics for Eq. 10.4 for Carbowax 20M	171
Table 10.9	Comparison of LFER coefficients for DEGS	171
Table 10.10	Regression Statistics for Eq. 10.4 for DEGS	172
Table 10.11	Solutes and descriptors used in Eq. 10.1 for squalane	174
Table 10.12	Solutes and descriptors used in Eq. 10.1 for Carbowax 20M	175

Table 10.13	Solutes and descriptors used in Eq. 10.1 for DEGS	176
Table 11.1	Table of the available solute descriptors	180
Table 11.2	Calculation of descriptors for 1,2-dicholoropropane	186
Table 11.3	Characteristic coefficients in Eq. (11.2) for partitions between water and solvents	187
Table 11.4	Characteristic coefficients in Eq. (11.1) for partitions between the gas phase and solvents	187
Table 11.5	List of θ values	188
Table 11.6	Determination of descriptors using log P, log K^W, log K and I values, when available, for some VOCs in table 11.12	189
Table 11.7	Determination of descriptors using log P, log K^W, log K and I values, when available, for some other VOCs	189
Table 11.8	Calculation of descriptors for phenylacetic acid	191
Table 11.9	Characteristic coefficients in Eq. (11.2) for partitions between water and solvents	192
Table 11.10	Characteristic coefficients in Eq. (11.1) for partitions between the gas phase and solvents	192
Table 11.11	Calculation of solvation descriptors by the method of 'leave-one-out'	193
Table 11.12	List of VOCs identified in indoor air	197
Table 12.1	Descriptors, as defined in Eq. 12.1, and values of log(1/NPT) for volatile organic compounds (VOCs)	208
Table 12.2	Regression coefficients in Eq. (12.1) for gas-solvent phase partitions at 298K	211
Table 12.3	Descriptors, as defined in Eq. 12.1, and values of log(1/ODT) for volatile organic compounds (VOCs)	213
Table 12.4	Calculated values of nasal pungency and odour detection thresholds for VOCs	221
Table 12.5	Relative significance of the descriptors, in percentage	229
Table 12.6	The effect of size on log(1/NPT) and log(1/ODT) values for some alcohols	229
Table 13.1	Deviations from Eq. (13.5) for various series	239

Table 13.2	Coefficients in equations for gas to solvent or phase transfer	242
Table 13.3	Compounds studied, their descriptors and values of log(1/ODT) and log(1/NPT)	247
Table 14.1	Nagata values of odour detection thresholds used to obtain Eq. 14.2	253
Table 14.2	Regression coefficients in Eq. (14.1) for gas-solvent phase partitions at 298K	259
Table 14.3	Compounds used to study the self-consistency of the Nagata data set	260
Table 14.4	Common compounds in the Nagata and Cometto-Muñiz and co-workers data set	263
Table 14.5	Four replicate values of log (1/ODT) with ODT in p.p.m.	267
Table 14.6	Analysis of Variance for the replicates of the data by Cometto-Muñiz and Cain	269
Table 14.7	Cometto-Muñiz and co-workers values of odour detection thresholds used in Eq. 14.6	270
Table 14.8	Hellman and Small values of odour detection thresholds used to obtain Eq. 14.9	273
Table 15.1	Characteristics of all subjects tested	281
Table 15.2	Value (±standard error) of parameters, and estimates of goodness of fit for odour functions modeled via the simplified sigmoid Eq. (15.2)	286
Table 15.3	Descriptors for the set of compounds used to build Eqs. (15.4) and (15.5)	289
Table 15.4	C and D as obtained from the psychometric curves, and the fitted (calculated) values of C and D from Eqs. (15.4) and (15.5). The average absolute error (AAE) has also been calculated for both parameters	294
Table 16.1	Characteristics of all subjects tested	300
Table 16.2	Value (±standard error) of parameters, and estimates of goodness of fit for nasal pungency functions modeled via the simplified sigmoid Eq. (16.2)	303
Table 16.3	Descriptors for the set of compounds used to build Eqs. (16.4) and (16.5)	305

Table 17.1 Dimensions (Å and $Å^3$) of the box enclosing the VOC for the 2-ketone series in vacuum 322

Table 17.2 Dimensions (Å and $Å^3$) of the box enclosing the VOC for the alkylbenzene series in vacuum 322

Table 17.3 Dimensions (Å and $Å^3$) of the box enclosing the VOC for the 2-ketone series in water 324

Table 17.4 Dimensions (Å and $Å^3$) of the box enclosing the VOC for the alkylbenzene series in water 324

List of figures

Figure 2.1	Diagrammatic representation of where the olfactory region is located	21
Figure 2.2	Diagrammatic representation of the olfactory system	23
Figure 2.3	Structure of the olfactory epithelium	24
Figure 2.4	Model of odorant signal transduction	27
Figure 2.5	Combinatorial Receptor Codes for Odorants. In this model, the receptors shown in colour are those that recognise the odorant on the left	31
Figure 2.6	Closely related odorants with different perceived odours are detected by different combinations of receptors	32
Figure 2.7	Transformations of odorant receptor inputs in the nervous system	33
Figure 2.8	Schematic diagrams showing the organization of odorant receptor inputs in the olfactory epithelium (OE), olfactory bulb (OB), and olfactory cortex (OC)	34
Figure 2.9	Representation of the regions within the nasal and oral cavities innervated by each of several cranial nerves	36
Figure 2.10	Schematic diagram of the branches of the trigeminal nerve that innervate the nasal, oral and ocular epithelia	36
Figure 2.11	Schematic representation of the brain stem	37
Figure 2.12	Procedures for presenting low concentrations of stimuli to subjects for assessment	39
Figure 2.13	Squeeze bottles with pop-out spouts	41
Figure 2.14	Various techniques for introducing odorant into a stream of non-odorous air	42
Figure 2.15	Triangle Odour Bag Method Source	46
Figure 3.1	Structure of the pre-ocular tear film of the human eye	58
Figure 3.2	(Left) Picture of a 1900 ml glass vessel adapted with an eyepiece for ocular stimulation. (Right) picture of a subject being tested for eye irritation	61
Figure 3.3	Psychometric function for butan-1-ol	62

Figure 4.1	The three steps of the cavity model of solvation	68
Figure 5.1	Schematic diagram of the principal components of a gas chromatograph	93
Figure 5.2	Example of a chromatogram	94
Figure 5.3	Structures of two hydrogen-bond stationary phases	111
Figure 6.1	Ionization equilibria of fluorescein	116
Figure 9.1	Ionization equilibria of fluorescein	142
Figure 9.2	Calculated concentration of the cation (□), lactone species (▽), monoanion (○), and dianion (△)	142
Figure 9.3	Spectrum of fluorescein at pH 3.28 in a cuvette with a width of 5 cm	147
Figure 9.4	3D Structure optimization	153
Figure 9.5	Spectrum of diacetylfluorescein in a cuvette exhibiting a width of 5 cm	156
Figure 9.6	Spectra of diacetylfluorescein in a 5 cm cuvette	157
Figure 9.7	Transesterification reaction	157
Figure 9.8	Spectra of diacetylfluorescein in a cuvette with a width of 5 cm	158
Figure 10.1	Structures of squalane, Carbowax 20M and diethylene glycol succinate	162
Figure 11.1	Plot of the calculated versus observed SP values for the systems in Table 11.2	187
Figure 12.1	Plot of log(1/NPT) predicted vs log(1/NPT) observed from Eq. 12.2	211
Figure 12.2	Histograms of the descriptors in the data set used to construct the nasal pungency threshold model	212
Figure 12.3	Histograms of the descriptors in the data set used to construct the odour detection threshold model	216
Figure 12.4	Residuals (observed-calculated values in Eq. (12.3)) against the L descriptor	217
Figure 12.5	Plot of log(1/ODT) predicted vs log(1/ODT) observed from Eq. 12.3	218
Figure 12.6	Scatterplot of log (1/ODT) against **L** for eight different homologous series of VOCs. Calculated line without the indicator	

variable (———). Calculated line with the indicator variable (– – – –). Outliers (○) 220

Figure 12.7 Histograms of the descriptors in Table 12.4 230

Figure 12.8 The two-stage mechanism of biological activity 232

Figure 13.1 A plot of log(1/ODT)obs against log(1/ODT) calculated on Eq. (13.5), showing the specific effects of aldehydes □ and acids ○. The regression line is for the selective compounds 238

Figure 13.2 A plot of log(1/ODT)obs against log(1ODT)calc on Eq (13.7) 241

Figure 13.3 The two-stage mechanism of biological activity of VOCs 243

Figure 13.4 A plot of the scores of PC2 against the scores of PC1. Points numbered as in Table 13.2: ● NPT, ▲ODT 244

Figure 14.1 Scatterplot of log (1/ODT) against **L** for eight different homologous series of VOCs. Calculated line without the indicator variable (———). Calculated line with the indicator variable (– – – –). Outliers (○) 257

Figure 14.2 Plot of observed values of log (1/ODT) by Nagata against the observed values of log (1/ODT) by Cometto-Muñiz and co-workers 264

Figure 14.3 Difference in the observed values of log (1/ODT) by Nagata and the observed values of log (1/ODT) by Cometto-Muñiz and co-workers. The average value of the differences and the confidence interval at 95% are included 265

Figure 14.4 Difference in the observed values of log (1/ODT) by Nagata and the observed values of log (1/ODT) by Cometto-Muñiz and co-workers after leaving out the outlier, 1-octene. The average value of the differences and the confidence interval at 95% are included 265

Figure 14.5 Difference in the observed values of log (1/ODT) by Nagata and the observed values of log (1/ODT) by Cometto-Muñiz and co-workers after leaving out the outlier, 1-octene, and subtracting the value 2.129 of the log (1/ODT) values by Nagata. The average value of the differences and the confidence interval at 95% are included 266

Figure 14.6 A plot of calculated values of log (1/ODT) on Eq. (14.7) against calculated values. ○ Nagata data; * Cometto-Muñiz and co-workers data; □ Hellman and Small data 275

Figure 14.7 A plot of the scores of PC2 against the scores of PC1 for the five Abraham descriptors in Eq. (14.7) for all 323 data points.

o Nagata data; * Cometto-Muñiz and co-workers data; Hellman and Small data 275

Figure 15.1 Odour psychometric functions modeled by Eq. (15.2) for chemicals and techniques listed in Table 15.2 284

Figure 15.2 Illustrating, against the background of all odour functions, how homologs gain olfactory potency, i.e., functions shift towards lower concentrations, with increasing carbon chain length 285

Figure 15.3 Comparison of the experimental (————) and calculated (– – – – –) odour psychometric functions for those chemicals used to obtain Eqs. (15.4) and (15.5) 291

Figure 15.4 Odour psychometric functions for those chemicals found as outliers, experimental (————) and calculated (– – – – –) 293

Figure 15.5 A plot of the experimental values of C against D 293

Figure 15.6 Structure of TXIB 295

Figure 15.7 Odour psychometric functions for TXIB and ethanol, experimental (———) and calculated (– – – – –) 296

Figure 16.1 Psychometric functions for nasal pungency modeled by Eq. (16.2) for chemicals and techniques listed in Table 16.2 302

Figure 16.2 Illustrating, against the background of all nasal pungency functions, how homologs gain nasal chemesthetic potency, i.e., functions shift towards lower concentrations, with increasing carbon chain length 303

Figure 16.3 Comparison of the experimental (————) and calculated (– – – – –) nasal pungency psychometric functions 306

Figure 16.4 Representation of the experimental values of C against D 307

Figure 17.1 Detectability (left *y*-axis) and confidence of detection (right *y*-axis) of the group (n=15) for the five chemicals presented to the eye at two flowrates 312

Figure 17.2 Individual ocular detectability of each chemical, averaged across the two flowrates of presentation, for the 15 subjects (dotted lines, each symbol represents one subject) 313

Figure 17.3 Showing: a) agreement in the ocular detectability of heptyl benzene and 2-tridecanone between the group in Experiment 1 (n=15) and that in Experiment 2 (n=21), both tested at room temperature, 23°C, and b) failure to enhance significantly the detectability of either chemical by increasing their vapor concentration via heating them to 37°C. 315

Figure 17.4 Individual ocular detectability of heptyl benzene and 2-tridecanone at room temperature, 23°C, and heated (H) to 37°C, for the 21 subjects (dotted lines, each symbol represents one subject) 316

Figure 17.5 Comparison between a group of unresponsive (n=13) and a group of responsive (n=8) subjects in terms of their ability to precipitate or increase the ocular detectability of the VOC with an increase in concentration (achieved by increasing the temperature of the VOC) 317

Figure 17.6 Relationship between various set of values for vapour concentration and carbon chain length for homologous 2-ketones 319

Figure 17.7 Analogous to Figure 6 but for homologous alkylbenzenes 319

Figure 17.8 A plot of the VOC box dimension, XYZ, against the number of carbon atoms for the 2-alkanone series in water 325

Figure 17.9 A plot of the VOC box dimension, XYZ, against the number of carbon atoms in the alkyl side chain for the alkylbenzene series in water 325

Ethical approval

All the experiments on humans were carried out by Enrique Cometto-Muñiz and co-workers, at the University of California, San Diego. The protocol for all these experiments on humans was approved by a committee of the Human Research Protections Program of the University of California, San Diego. All subjects gave written informed consent on forms approved by the committee.

Abbreviations

A	LFER constant
AAE	Absolute Average Error
AE	Average Error
AFC	Alternative Forced-Choice
ANOVA	Analysis of Variance
B	LFER constant
BM	Biological Monitoring
BRI	Building Related Illness
CCS	Common Chemical Sense
E	LFER constant
ECD	Electron Capture Detector
EIT	Eye Irritation Threshold
EPA	Environmental Protection Agency
ETS	Environmental Tobacco Smoke
F	The Fisher F-statistic
FID	Flame Injector Detector
GC	Gas Chromatography
GLC	Gas Liquid Chromatography
GV	Glass Vessel
HBA	Hydrogen Bond Acceptor
HBD	Hydrogen Bond Donor
HPLC	High Performance Liquid Chromatography
IAP	Indoor Air Pollutants
IAPI	Indoor Air Pollution Index
IETS	Inelastic Electron Tunnelling Spectroscope
IUPAC	International Union of Pure and Applied Chemistry
Ks	Gas to solvent partition coefficient
Kw	Gas to water partition coefficient
L	LFER constant
LFER	Linear Free Energy Relationship
MCS	Multiple Chemical Sensitivity

MLRA	Multiple Linear Regression Analysis
N	Number of data points
NPT	Nasal Pungency Threshold
OB	Olfactory Bulb
OBP	Odorant Binding Proteins
OC	Olfactory Cortex
ODT	Odour Detection Threshold
OE	Olfactory Epithelium
OR	Odorant Receptors
ORN	Olfactory Receptor Neurons
OSN	Olfactory Sensory Neurons
P	Water to solvent partition coefficient
PCA	Principal Component Analysis
QSAR	Quantitative Structure Activity Relationship
R	Correlation coefficient
RMSE	Root mean squared error
S	LFER constant
SAR	Structure Activity Relationship
SB	Squeeze Bottle
SBS	Sick Building Syndrome
SD	Standard Deviation
SMILES	Simplified Molecular Input Line Entry System
SV	Saturated Vapour
TEAM	Total Exposure Assessment Measurement
TVOC	Total Volatile Organic Compounds
V	LFER constant
VDD	Vapour Delivery Device
VOC	Volatile Organic Compounds
WHO	World Health Organization

Acknowledgements

My sincere thanks go to Prof Michael Abraham for his help over the last three and half years with this project. His knowledge and expertise in the area of physical chemistry has helped me to accomplish my PhD. I am extremely grateful for his unfailing encouragement, patience and guidance.

I would like to thank particularly my research group, Elizabeth Hills, Dr Adam Ibrahim and Dr Yuan Zhao, who have provided me with encouragement and great friendship over the years. A special thank to Javier Gil-Lostes for his help and support during the last couple of years. I wouldn't have finished this without your help. Thanks also to Dr Joelle Gola and Dr Juliette Osei-Owusu, who spent time in the laboratory giving me their knowledge on the use of the chromatographs.

I extended my gratitude to Dr Enrique Cometto-Muñiz and Dr William Cain for all the chemosensory threshold values included in this work, and for providing me with their expertise and advice on various issues.

My thanks also go to the University of California, San Diego, for the funding of this project and giving me the opportunity to study for the PhD.

In addition, I would like to thank all those in the chemistry department who have helped me in any way with my work or in making my time spent at University College London a happy one. In particular, Jill Maxwell and Jorge Sánchez-Galiani for their tireless understanding and for always being there when I needed them. Thank you!

Importantly, I would like to thank my family for their support during all these years of my PhD. My mum, to whom this thesis is dedicated, for having shown never-ending encouragement, love and support. A big thank you also to my brother, Gabi, I would not have got so far without his help and support. I would also like to thank my cousin Diana and her family, who welcomed me when I arrived to London, and made all the time I spent with them very happy.

Finally, my heart goes to my special fiancé (Miss Consuelo Verdejo) for her constant love and support during my PhD. Thanks for waiting for me to be back in Spain.

¡Gracias!

Chapter 1 An introduction to Indoor Air Pollution

1.0. INTRODUCTION

Until recently, the health effects of indoor air pollution have received relatively little attention from the scientific community. In the early 1970s, problems with indoor air quality in residences and non-industrial workplace were occasionally investigated, but the level of interest was low [1]. Numerous studies suggest that many individuals perceive the risks from poor quality outdoor air as being substantially higher than those from indoor contamination [2]. These perceptions exist despite the fact that, in developed societies, the majority of people pass most of their time indoors. For instance, a recent investigation of time budgets amongst US residents found that, on average, individuals spent 88% of their day inside buildings, and 7% in a vehicle. Only 5% of participants' time was actually spent outside [3]. Whilst this does not by itself means that indoor exposures will produce more harmful health effects, the fact that indoor concentrations of many pollutants are often higher than those typically found outside, lends weight to the serious nature of indoor air pollution.

A preoccupation with the relationship between outdoor air quality and health is understandable. Particularly in more economically developed and rapidly developing countries, recent and unprecedented changes in lifestyles and environmental quality have meant that an increasing number of people are exposed to the contaminations of urban air [4]. There is certainly evidence that, especially amongst more vulnerable members of society, outdoor pollutants may pose a real health risk [5].

Nowadays, concern over the health effects of poor quality indoor air is increasing [6]. Despite the fact that the vast majority of buildings exhibit no immediately apparent problems, a wide spectrum of symptoms and illnesses are attributed to non-industrial indoor air pollution [7].

Indoor air pollutants (IAPs) vary from heavy metals to volatile organic compounds, VOCs; it is the latter that is emphasised in this chapter. VOCs have a wide range of physical and chemical characteristics. Chemical groups typically include hydrocarbons, halogenated hydrocarbons, aromatic hydrocarbons, alcohols, esters, ketones and aldehydes. They are defined as compounds that have melting points below

room temperature and boiling points from 50-100 to 240-260 °C [8].

VOCs are important air pollutants for two reasons. Firstly, some VOCs are together with nitrogen oxides (NO_x) precursors required for the photochemical production of ozone (O_3) and other components of photochemical smog [9]. Secondly, VOCs include species that have individually severe or chronic health effects in humans [10].

The number of identified VOCs present in ambient air has risen steadily in recent years, from 250 to more than 900 in 1989, to well over 1,000 currently [11]. As the number as well as the diversity of VOCs in the air can be so large, the concept of 'Total Volatile Organic Compounds' (TVOCs) has been used in many publications rather than individual values to report on concentrations of VOCs in the air. It has been used both for reporting exposures, that is, as indicator of air quality, and as a predictor of the probability of health and comfort effects. However, no documented background exists for the use of the TVOC indicator in relation to health and discomfort other than sensory irritation (e.g. not for cancer, allergy, and neurological effects).

1.1. SOURCES OF INDOOR AIR POLLUTANTS

Indoor air pollutants emanate from a range of sources. They are emitted by the fabric of buildings, and may also be a by-product of the activities that are undertaken within them. Sources can be broadly classified as being associated with the activities of building occupants and other biological sources, the combustion of substances for heating or fuel, and emissions from building materials. For some contaminants, infiltration from outside, either through water, air, or soil, can also be a significant source [12]. In fact, outdoor sources may be the main contributor to indoor concentrations of a number of non-biological pollutants commonly found in indoor air. This is especially the case for contamination in buildings situated in urban areas and close to industrial zones or streets with heavy traffic. The major outdoor sources of important indoor air pollutants are given in Table 1.1.

Heating and cooking are often essential indoor activities but may produce smoke and gases that present a problem of disposal. Combustion is another source for a wide range of pollutants. Cigarette smoking is an important source of indoor combustion related pollution. Tobacco smoke is an aerosol containing several thousand substances that occur as particles, vapours and gases. Materials that comprise the fabric of

buildings are another important source of non-biological pollutants. A large number of chemical compounds which are found in indoor air originate from paints, varnishes, solvents and wood preservatives used in buildings [13]. Furthermore, deteriorating materials that comprise the fabric of a building can become friable and release contaminants in to the air. Pollutants from these sources are often difficult to quantify because they are present in relatively low concentrations, and their sources are diffuse [14].

Table 1.1
Outdoor sources of major indoor air pollutants

Pollutant	Percentage of emissions associated with industry[a]	Percentage of emission associated with transport[a]
Benzene	32	65
Carbon Monoxide (CO)	3	90
Lead (Pb)	31	60
Oxides of nitrogen (NO_x)	38	49
Particulates (PM_{10})	56	25
Sulphur dioxide (SO_2)	90	2
VOCs	52	34
Ozone (O_3)	Arises from atmospheric chemical reactions	

[a] Figures based on UK estimates (DOE, 1997)

Another group of pollutants come from biological sources. Biological agents such as fungi, bacteria and viruses present in indoor air environments can cause disease through a topic mechanism, infectious, processes, or direct toxic effects. Dust-mite, cat, dog, rodent, cockroach, fungi are common indoor allergic agents [12]. House dust in carpets, on sofas, and in air ducts is a major source of a range biological allergens [15].

1.1.1 Volatile Organic Compounds

So far, a large amount of source and pollutants have been mentioned. However, the aim of this section is to focus the attention on volatile organic compounds which will be object of study in this project.

Main anthropogenic sources of VOCs include industrial processes, fossil fuel combustion in transportation and electricity generation sectors, household products, landfills and waste treatment plants. In many areas gasoline vapour and motor vehicle exhaust are the major sources for most VOCs. Overall, vehicle-related VOC emissions

are ubiquitous in urban areas. In addition, solvents from architectural and industrial sources are common but highly variable in most areas, while petrochemical production and oil refining are more specific to certain urban settings with theses activities [16].

There are also natural sources that generate VOCs in urban areas through direct emissions and through photochemical oxidation of naturally emitted hydrocarbon precursors [17]. Plants synthesize many organic compounds such as ethene, aldehydes, ketones, alcohols, isoprene and terpenes as an integral part of their biochemistry. Deciduous trees have found to be mainly isoprene emitters, while conifers favor monoterpenes [18].

Primary sources of VOCs in indoor environments include outdoor air (penetration from outdoors to indoors) as well as sources such as tobacco smoke, fuel combustion, building materials, furnishing, furniture and carpet adhesives, cleaning agents, ventilation system, cosmetics and the occupants themselves. Table 1.2 lists sources of common VOCs found indoors.

Tobacco smoke is an aerosol containing several thousand substances that are distributed as particles, vapours, and gases (VOCs) [19]. Environmental tobacco smoke (ETS) is possibly one of the most important indoor air pollutants in homes and office [14], this is because a substantial proportion of the population are exposed to high concentrations of ETS on a regular basis. Exposure to environmental tabacco smoke is a source of particular concern for both smoker and non-smoker. Studies have identified over 4000 compounds in smoke generated from burning tobacco, many of which are known carcinogens, mutagens and teratogens [20]. Acute respiratory effects of ETS exposure on adults have been demonstrated in experimental studies using exposure chamber. The most common acute effects of ETS exposure were irritation of nose, throat and airways [21].

Table 1.2
Sources of common volatile organic compounds in indoor air

Sources	Examples of typical contaminants
Consumer and commercial products	n-decane, branched alkanes, toluene, xylenes, methylene chloride, alcohols, acetone, methyl ethyl ketone, formaldehyde, alkyl ethoxilate, glycol ether, limonene, alpha-pinene
Potable water	1,1,1-trichloroethane, chloroform
Combustion appliances	Propane, butane, isobutene, acetaldehyde, acrolein

Table 1.2 *(Continued)*

Sources	Examples of typical contaminants
Furnishings and clothing	styrene, brominated aromatics, vinyl chloride, formaldehyde, ethers, eters
Adhesives	hexane, heptane, aromatic hydrocarbons, halogenated hydrocarbons, alcohols, amines, acetone, methyl ethyl ketone, vinyl acetate
Paints and associated supplies	n-hexane, n-heptane, toluene, methylene chloride, propylene chloride, alcohols, methyl ethyl ketone, ethyl acetate, methyl ether, ethyl ether, butyl ether
Building materials	n-decane, n-dodecane, toluene, styrene ethylbenzene, vinyl chloride, formaldehyde, acetone, butanone, urethane, ethylacetate

Adapted from Ref. 22

1.2. ASSESSMENT OF EXPOSURE

Assessment of exposure in humans refers to the analysis of various processes that lead to human contact with pollutants after release in the environment [23]. The term exposure refers to the length of contact with the pollutant during a specified period of time. Assessment of exposure is a science on its own. Here, attention is mainly drawn to methods used to determine pollutant concentration in indoor environments.

A wide range of sampling and analytical methods has been applied to determine the nature and concentration of VOCs [24,25]. The most common methods are based on collection using absorbents, e.g. Tenax GR, contained in a sampling tube or badge. The absorbent can be thermally desorbed or an extraction with solvent can be carried out instead, and the VOCs usually are determined by gas chromatography coupled with various detection systems such as FID or ECD. Other absorbents are available and are chosen according to the polarity of the VOCs investigated. The sampler can be used in either an active or a passive mode. To obtain sufficient sensitivity for analysis the exposure period of the passive sampler is typically between one to four weeks whereas for active sampling this can be achieved with sampling times of one hour. Other sampling methods are grab sampling, condensation, or liquid or solid phase extraction. Solid-phase or liquid extraction is commonly used for less volatile compounds.

Discomfort experienced in poor indoor air quality environment, is caused by simultaneous presence of individuals and air pollutants or in other words, individual's

exposure to air pollutants. The individual or personal exposure level is best assessed by measuring an individual's contact with pollutants using personal monitoring or biological monitoring [26].

1.2.1 Personal Monitoring

Measurements from personal monitoring indicate the level of external exposure [24]. They are carried out by means of small devices, which sample IAPs, placed on the individuals. For instance, bubblers, vapour adsorption tubes and passive samplers are widely used in personal monitoring to measure the concentration of airborne volatile chemical in the region of the mouth [27]. These measurements can be carried out in real environment or in simulated exposure chamber as explained later.

1.2.2 Biological Monitoring

According to the International Union of Pure and Applied Chemistry (IUPAC) [23], biological monitoring (BM) is a "systematic continuous or repeated measurement and assessment of workplace agents or their metabolites either in tissues secreta, excreta or any combination of these to evaluate exposure or health risk compared to appropriate reference". In other words, BM allows one to assess the integrated exposure by different routes, including ingestion, inhalation, dermal absorption, blood, exhaled air and urine.

Biological marker or biomarkers for exposure is an endogeneous substance or its metabolite or the product of interaction between a xenobiotic and some target molecule or cell that is measured in a compartment with an organism [23]. Biomarkers are focused on the amount of the pollutant penetrating to the organism. Several biomarkers are relevant to indoor air pollution, for instance, nicotine has been advocated as an appropriate biomarker for estimating quantitative exposure to ETS because it is a direct component of cigarette smoke that is not present in air from other combustion sources [28]. Other examples are the use of carboxyhaemoglobin level in blood to characterise exposure to CO, and the use of the presence of VOCs in exhaled air breath to mark these compounds [27].

1.2.3 Controlled Exposure Chamber

Assessment of exposure is mainly carried out according to the above methods. These techniques can be used in real environments e.g. offices, homes, vehicles, or in controlled or simulated exposure chambers. Controlled exposure chambers have been devised to study single or multiple pollutants in relative 'pure' form, without potential interference of other materials. These exposure systems vary from small volume, personalised mouth-piece or face mask configurations to large chambers, which are often used to characterise pollutants emitted from wallpaper, carpet, and other materials. Beside, small devices are more appropriate for controlled human exposure studies, which are carried out to evaluate human response to pollutants. The main drawback is that true simulation of the 'real' environment is impractical [26].

1.3. CONCENTRATION OF INDOOR AIR POLLUTANTS

On one hand, as mentioned, the outdoor sources can be the main contributor to indoor concentrations. The factors that determine the contribution of outdoor pollution to indoor air quality include the type of ventilation in use (natural or forced), the ventilation rate (air changes per hour), and the nature of contaminants in question [12]. On the other hand, the emissions of IAPs from materials will vary with the age of the source, the type of pollution transfer process (evaporation, desorption, or diffusion), and the impact of environmental factors such as temperature, humidity, air changes, and air velocity. Levels of indoor air pollutants are also dependent on human's activity. Because these factors are constantly changing, the concentration will also change with time and how dramatic that change is will depend on the circumstances. An evaluation of the change of indoor concentration of VOCs in apartments and a comparison with outdoor concentration levels have been carried out by Rehwagen et al. [29] and by Ohura *et al.* In this one, also the seasonal variability (winter/summer) was taken into account [30]. The impact of temperature and humidity to study emission from materials has been evaluated by Fang et al. [31].

It has been estimated that in a typically non-industrial indoor environments 50-300 different VOCs are continuously present in the air [32]. The United States Environmental Protection Agency (EPA) has carried out a number of studies to determine levels and also exposure to IAPs in several urban, non-urban and industrial

and non-industrial areas in various places in the US. One of the most comprehensive studies designed to determine the exposure of individuals to IAPs within their homes was the Total Exposure Assessment Measurement (TEAM) study [24,33]. This study pointed out that personal exposures to VOCs were two to five times higher than outdoor concentration, even though the outdoor concentration were measured in heavily polluted areas such northern New Jersey and Los Angeles. Further, much of the difference is attributable to exposure to indoor sources, such as environmental tobacco smoke [33].

The US EPA conducted a study about indoor and outdoor air concentrations of twenty-seven hazardous air pollutants [33]. Indoor and outdoor exposures (i.e. concentrations breathed multiplied by duration of time breathed) were estimated from data of the literature. The results emerging from this study showed that the indoor air concentration of these hazardous compounds are generally one to five times outdoor concentration, and indoor exposures are ten to fifty times outdoor exposures as displayed in Table 1.3.

Table 1.3
Summary of Indoor and Outdoor Air concentrations and Exposures

	Typical Concentration [a]			Typical Daily Exposure [b]		
HAP	Indoor	Outdoor	I/O Ratio	Indoor	Outdoor	I/O Ratio
Acetaldehyde [c]	<10	<3	*[5]*	<216	<7.2	*[50]*
Benzene [c]	5	5	*[2]*	108	12	*[20]*
Captan	<0.001	Na	10	<0.02	<0.0002	~100
Carbon tetrachloride[c]	<5	1	*[2]*	<108	2.4	*[20]*
Chlordane	0.2	0.01	20	4.32	0.024	200
Chloroform[c]	1	0.2	*[5]*	21.6	0.48	*[50]*
Cumene	1	0.2	5	21.6	0.48	50
2,4-D (salts, esters)	0.001	0.00003	30	0.0216	0.000072	300
DDE	0.0005	<0.002	>0.2	0.0108	<0.005	>2
Dichlorvos	0.05	0.001	50	1.08	0.0024	400
Ethylbenzene	5	1	*[3]*	108	2.4	*[30]*
Formaldehyde	50	4	10	1080	9.6	100
Formaldehyde[c]	0.1	0.002	50	2.16	0.0048	400
Hexachlorobenzene	0.0005	0.0001	5	0.0108	0.00024	50

Table 1.3 *(Continued)*

	Typical Concentration [a]			Typical Daily Exposure [b]		
HAP	Indoor	Outdoor	I/O Ratio	Indoor	Outdoor	I/O Ratio
Hexane	5	4	*[10]*	108	9.6	*[90]*
Methoxychloride	0.0001	0.00003	3	0.00216	0.000072	30
Methyl ethyl ketone	10	<1	*[4]*	216	<2.4	*[40]*
Methylene chloride	10	1	*[10]*	216	2.4	*[90]*
Naphthalene	1	1	*[4]*	21	2.4	*[40]*
Para-dichlorobenzene	1	<0.05	*[5]*	21.6	<0.12	*[50]*
Propoxur	0.1	0.01	10	2.16	0.024	90
Radionuclides[d]	2	0.1	20	43.2	0.24	200
Styrene [c]	2	0.6	*[5]*	43.2	1.44	*[50]*
Tetrachloroethylene[c]	5	2	*[3]*	108	4.8	*[30]*
Toluene	20	5	*[5]*	432	12	*[50]*
Trichloroethylene[c]	5	0.5	*[5]*	108	1.2	*[50]*
Xylenes (o+m+p)	15	10	*[2]*	324	24	*[20]*

[a]Indoor and outdoor concentrations in $\mu g \cdot m^{-3}$, Typical values from the literature for U.S. locations. I/O ratios based on typical concentrations *[or reported ratios as indicated by values in italics and brackets].* [b]Exposure, in $\mu g \cdot m^{-3} \cdot h^{-1}$, based on assumption that the typical person spends 90% of the typical day indoors. [c]Urban Air Toxic substances, [d]Radionuclides/Rn in $pCi \cdot .dm^{-3}$.

Note$_1$: DDE includes 2,2-Bis(4-chlorophenyl)-1,1-dichloroethylene and 2-(2-Chlorophenyl)-2-(4-chlorophenyl)-1,1-dichloroethene

Note$_2$: The chemical name of 2,4-D is 2,4-Dichlorophenoxyacetic acid

Adapted from Ref. 33

A series of studies on personal exposure to environmental tobacco smoke (ETS) has been carried out in several countries [24,34]. Some results emerging from these studies showed that persons with a smoking partner had mean exposure of 219 $\mu g \cdot m^{-3}$, compared to about 170 $\mu g \cdot m^{-3}$ for persons without a smoking partner, a difference of about 49 $\mu g \cdot m^{-3}$ for those without [24]. Studies have also shown that where smoking rates are high and ventilation minimal there is a clear contribution to formaldehyde concentrations from ETS of the order of a few tens of $\mu g \cdot m^{-3}$. Concentrations of aliphatic and monocyclic aromatic hydrocarbons can rise from 2-20 $\mu g \cdot m^{-3}$ to 50-200 $\mu g \cdot m^{-3}$. Detailed studies in the USA of residential exposures suggest that non-smoking

households experience an exposure to approximately 7 $\mu g \cdot m^{-3}$ of benzene, while households with a smoker are exposed to approximately 11 $\mu g \cdot m^{-3}$. Excess exposure to styrene and xylenes range from 0.5 to 1.5 $\mu g \cdot m^{-3}$ [34].

Because of the many compounds present simultaneously in office buildings or houses, which make detailed measuring programs time consuming and expensive, and because of the difficult to relate a symptom with a particular compound, the TVOC-indicator was born. The total volatile organic compounds (TVOC mass/m^3) was suggested as an indicator of exposure to a mixture of VOCs and of its potency to cause effects [35-39]. On the other hand, a Nordick working group has evaluated the use of TVOC and concluded, "the group cannot recommend the present use of TVOC as a risk indicator for health effects and discomfort problems in building" [40]. Other researchers have reached a similar conclusion [41,42]. As a result, TVOC has limited use and must be used with precaution, as an indicator to relate exposure characterization and source identification just for VOCs, and as a screening tool for exposure evaluation but only for sensory effects [39].

1.4. INDOOR AIR POLLUTANTS AND HEALTH

Exposure to indoor air pollution is a potentially serious public health problem in a wide variety of non-industrial settings, for instance, residences, offices, schools, and vehicles [43]. Exposure to airborne toxicants indoor can potentially lead to a variety of adverse health outcomes, depending on variables such as inherent toxicity of the agent or agents in question, exposure magnitude, duration, frequency and route, and intra-individual and interindividual variability in biological susceptibility to the effects of exposures.

Currently, indoor air pollution is ranked by the US EPA and the Centers for Disease Control and Prevention in the top five environmental risks [44].

Immediate effects reported by building occupants have been classed in various categories whose most widely referred to are:

- Multiple Chemical Sensitivity (MCS)
- Building-Related Illness (BRI)
- Sick Building Syndrome (SBS)

1.4.1 Multiple Chemical Sensitivity

In most cases individuals describe unexplainable symptoms and signs of unexplainable dysfunctions which may or may not be related to their environment. Exposures to low levels of chemicals have been suggested as causative agents for some of these health effects and have given rise to a public health concern most often called multiple chemical sensitivity (MCS). The history of MCS has been summarized by Shorter [45].

Most authors seem to refer to an intolerance or hyperresponsiveness to exposure to low levels of chemicals at home, during work, or in the course of their life activities. Since MCS seem to affect few individual occupants with unusual high sensitivity, MCS must be different from sick building syndrome (SBS), which affects a large fraction of building occupants [37].

1.4.2 Building-Related Illness

Building-Related Illness (BRI) has a known aetiology with specific symptoms and laboratory findings. The most common diseases are characterized by flu-like symptoms or respiratory distress and include hypersensitivity pneumonitis, humidifier fever, and bacterial pneumonia, e.g. Legionnaire's disease; Pontiac fever. BRI also includes any toxic reactions to chemicals used within the building, or derived from fungi growing within a building. The incidence of these disorders is apparently unrelated to building characteristics like VOC levels [46].

1.4.3 Sick Building Syndrome

Sick building syndrome is normally distinguished from the term 'building-related illness'. Whilst BRI has a known aetiology, as mentioned, many studies of SBS have failed to determine any specific factor that may explain the health complaints. SBS is a situation that occurs when indoor air quality deteriorates and a statistically significant number of individuals in unique portions of a building complain about the following [47]:

- nose and throat: dryness, stuffiness, irritation of the mucous membrane, runny nose, sneezing, odour, taste sensations

- eyes: dryness, itching, irritation, watery eyes
- skin: dryness and irritation, rash itching, eczema
- general: tight-chest, abnormal fatigue, malaise, headache, feeling heavy-headed, lethargy, sluggishness, dizziness, nausea

All these symptoms are common in the general population; the distinguishing feature which makes them part of the sick building syndrome is their temporal relation with work in a particular building. They usually begin within 15 minutes of entering the building and all except skin symptoms should improve within a few hours of leaving a problem building; dryness of the skin may take a few days to improve [48]. The syndrome seems to be a normal reaction of a normal population to unfavourable indoor climate. Thus, the WHO group suggested the possibility that the SBS symptoms have a common causality and mechanism [49].

In Japan, SBS symptoms associated with indoor air VOCs in new or newly remodelled houses have been increasingly highlighted, and they are called 'Sick House Syndrome' [50].

One of the factors that seems play an important role in SBS is building ventilation systems as mentioned elsewhere [51,52]. However, there are other findings that show no significant difference in symptom prevalence amongst the occupants of buildings with ventilation rates between 13 and 42 cubic feet per minute per person ($cfm \cdot p^{-1}$)[53].

Despite the voluminous research undertaken over the last two decades to investigate the reason for SBS outbreaks, it has been estimated that no specific cause has been identified in over 75% of cases [54]. This may, in part, be due to methodological considerations. One problem is that currently available pollutant sampling techniques are costly and suffer from some technical limitations, and another problem is the lack of a standardised system for the reporting and diagnosis of symptoms [12]. An attempt to overcome this last situation was carried out by Mølhave and co-workers [34,35,39,55,56], as they proposed the use of TVOC as an indicator of indoor air quality and they tried to link this indicator with health symptoms. However, these attempts to associate occupant symptoms and TVOC levels report inconsistent findings, as mentioned in section 1.3. Several suggestions have been made to use different indoor air pollution indexes, most of them with few applications and without relationships between these indexes and occupant symptoms. In a recent work by

Sofuoglu and Moschadreas, a new indoor air pollution index (IAPI) was formulated [57]. In this case, the index includes a bigger number of pollutants: bacteria, carbon monoxide, carbon dioxide, formaldehyde (HCHO), fungi, $PM_{2.5}$, PM_{10}, radon and TVOC, in a linear function to estimate a single number, the IAPI. Furthermore, this index produces a good link with symptoms of office buildings occupants, which was the aim of developing this index.

1.4.4 Volatile organic compounds and health effects

As mentioned, VOCs are frequent air pollutants in non-industrial environments, in fact, normally 50 to 300 VOCs are found in air samples from most of these non-industrial environments [55]. Exposure to VOCs can result in both acute and chronic health effects [12], they can cause or contribute to a wide range of health effects from non-specific sensory response to specific toxicity target organs. Some selected volatile organic compounds and their health effects are listed in Table 1.4.

Exposure to high concentrations of many VOCs have been associated with symptoms of sick building syndrome, many of them are narcotics, and can depress the central nervous system [22]. Exposures can also lead to irritation of the eyes and respiratory tract, and cause sensitisation reactions involving the eyes, skin, and lugs.

Table 1.4
Health Effects of selected Volatile Organic Compounds

VOC	Health Effects
Benzene	Carcinogen; respiratory tract irritant
Xylenes	Narcotic; irritant; affects heart, liver, kidney, and nervous system
Toluene	Narcotic; possible cause of anaemia
Styrene	Narcotic; affects control of nervous system; probable human carcinogen
Toluene diisocyanate	Sensitizer; probable human carcinogen
Trichloroethane	Affects central nervous system
Ethyl benzene	Severe irritation of eyes and respiratory tract; affects central nervous system
Dichloromethane	Narcotic; affects control of nervous system; probable human carcinogen
1,4-Dichlorobenzene	Narcotic; eyes and respiratory tract irritant; affects heart, liver, kidney, and nervous system

Table 1.4 *(Continued)*

VOC	Health Effects
Benzyl chloride	Central nervous system irritant depressant, affects liver and kidney; eye and respiratory tract irritant
2-Butanone	Irritant; central nervous system depressant
Petroleum distillate	Affects central nervous system, liver and kidneys
4-Phenylcyclohexene	Eye and respiratory tract irritant; central nervous system effects

Adapted from Ref. 58

At extreme concentrations, some VOCs may result in impaired neurobehavioral function [59]. At concentration as high as 188 $\mu g \cdot m^{-3}$, VOCs such as toluene may cause symptoms of lethargy, dizziness, and confusion. These may progress to coma, convulsions and possibly death at levels in excess of 35.000 $\mu g \cdot m^{-3}$ [60]. However, these concentrations have never been recorded in the non-industrial environment.

The question is still open whether concentrations of VOCs in the non-industrial indoor air are sufficiently below the threshold for comfort-reducing irritation and other symptoms included in the sick building syndrome. To test if low exposure levels of VOCs may cause discomfort, five controlled exposure experiments in climate chambers were summarized by Mølhave [61], see Table 1.5.

Table 1.5
Effects of Human Health and Well-being Caused by Exposure to VOC in Indoor Climate

	VOC Exposure		
Effect	0-2.9	3-25	> 25
Sensory irritation	1/1/4	5/0/1	2/0/4
Olfaction	2/0/4	5/0/1	2/0/4
Toxic irritation (tissue changes)	1/1/4	5/0/1	2/0/4
Weak neurological	1/1/4	5/0/1	2/0/4
Lung and lower air	0/0/6	0/0/6	1/0/5
Allergy, etc.	0/0/6	0/0/6	0/0/6
Systematic and organ	0/0/6	0/0/6	0/0/6

Note: Six investigations (5 exposures and a review of 13 field investigations). The numbers show the numbers of these studies indicating positive or negative effects or the effect not investigated (positive/negative/not investigated).

Adapted from Ref. 61

The five exposure experiments provide no consistent information about lung function or allergic or systematic effects at any exposure levels and, except for olfaction, no effects at exposure levels below 3 $mg \cdot m^{-3}$. At present the response of the olfactory sense seems to be the most sensitive indicator of VOC exposure.

Investigations of sensory irritation, olfaction, skin irritation, and weak neurological effects caused by exposure levels of 3 $mg \cdot m^{-3}$ and higher have indicated effects in all experiments dealing with such exposure levels [37].

The intensity of each of the symptoms may be modified by additional factors such as age, smoking habits, and gender. Each subject may react differently to the mixed exposure and exhibit only a few of the spectrum of symptoms observed in the exposed population [37].

To discuss the relationship between VOCs and the different symptoms included in the sick building syndrome must be considered not only the VOC exposure levels but also the levels of other contributing exposures and the reactions between them. In fact, a recent theory proposes that the products of chemical reactions involving VOCs may be more important than direct exposure to the VOCs themselves [62].

1.5. REFERENCES

1. J. A. Stolwijk, Ann. N.Y. Acad. Sci., 641 (1992) 56
2. London Health Education Authority (LHEA), *What people think about air pollution, their health in general, and asthma in particular*, 1997
3. J. Robinson, W. C. Nelson, National Human Activity Pattern Survey Data Base. United States Environmental Protection Agency, Research Triangle Park, N. C., 1995
4. F. W. Lipfert, Risk Analysis 17 (1997) 137
5. M. L. Burr, Arch. Dis. Child. 72 (1995) 377
6. Nguyen Thi Kim Oanh and Yung-Tse Hung in *Advance Air and Noise Pollution Control*, L. K. Wang, N. C. Pereira, Yung-Tse Hung (eds). Humana Press Inc. Totowa, New Jersey, 2005: pp 237-272
7. C. A. Redlich, J. Sparer, M. R. Cullen, Lancet 349 (1997) 1013
8. WHO (World Health Organisation). *Indoor Air Quality: Organic Pollutants*. European Reports and Studies No. 111. Copenhagen, Denmark, 1989
9. S. E. Manahan, *Environmental Chemistry*, 6th Edition, Lewis Publishers. US, 1994

10. WHO. Guidelines for Air Quality. Geneva, Switzerland, 2000
11. E. W. Miller, R. M. Miller, *Indoor Air Pollution: a reference handbook*, ABC Clio, Santa-Barbara, 1998
12. A. P. Jones, Atmos. Environ. 33 (1999) 4535
13. European Commission Joint Research Centre, *Total volatile organic compounds (TVOCs) in indoor air quality investigations*, European Commission, Luxembourg, 1997
14. H. U. Wanner in *Sources of pollutants in indoor air*. IARC Scientific Publications, 1993, 109, 19-30
15. R. G. Lewis, R. C. Fortmann, D. E. Camann, Arch. Environ. Cont. Toxic. 26 (1994) 37
16. J. G. Wason, J. C. Chow, E. M. Fujita, Atmos. Environ. 35 (2001) 1567
17. J. H. Seinfeld, S.N. Pandis., *Atmospheric Chemistry and Physics: from Air Pollution to Climate Change*, John Wiley & Sons Inc, New York, 1998
18. J. Colls, *Air Pollution - an introduction,* E&FN Spon, an Imprint of Chapman & Hall. London, United Kingdom, 1997
19. D. R. Gold, Clinics in Chest Medicine 13 (1992) 215
20. D. Roberts, Recent Adv. Tobacco Sci. 14 (1988) 49
21. J. Tetradaniel, P. Boffetta, R. Saracci, A. Hirsch, Eur. Respir. J. 7 (1994) 173
22. M. Maroni, B. Seifert, T. Lindvall. *Indoor Air Quality, A Comprehensive Reference Book,* Vol. 3, Elsevier Science. Amsterdam, The Netherlands, 1995
23. J. H. Duffus. Pure Appl. Chem. 65 (1993) 2003
24. L. Wallace, *Total Exposure Assessment Methodology (TEAM) Study: Summary and Analysis*, Vol.1, US EPA, 1987
25. R. E. Clement, P. W. Yang, Anal. Chem. 71 (1999) 257
26. H. Janturen, J. J. K. Jaakola, M. Krzyzanowski, *Assessment of Exposure to Indoor Air Pollutant*, WHO regional Publication European Series, Copenhagen, 78 (1997)
27. L. J. Folinsbee, C. S. Kim, H. R. Hehrl, J. D. Prah, R. B. Devlin in *Handbook of human toxicology*, E. J. Massaro (ed), CRC Press, New York, US, 1997
28. A. C. Collier, C. A. Pristos, Chemico-Biological Interactions 146 (2003) 211
29. M. Rehwagen, U. Schlink, O. Herbath, Indoor Air 13 (2003) 283
30. T. Ohura, T. Amagai, Y. Senga, M. Fusaya, Sci. Total Environ. 366 (2006) 485

31. L. Fang, G. Clausen, P. O. Fanger, Indoor Air 9 (1999) 193
32. L. Mølhave, *Volatile Organic Compounds, Indoor Air Quality and Health.* In: Proceedings of Indoor Air '90, Vol. 5. International Conference on Indoor Air Quality and Climate. Ottawa, Canada, 1990: pp. 15-33.
33. US EPA (United States Environmental Protection Agency), *Inside IAQ*, EPA-600-N-98-002 (1998) 1. US Environmental Protection Agency, Washington, D.C.
34. L. Sheldon, H. Zeton, J. Sickes, C. Eaton, T. Hartwell, *Indoor Air Quality in Public Buildings*, Vol. II, Research Triangle NC: EPA/600/6-88/009b, 1988b.
35. L. Mølhave, ASHARE Trans. Paper 2594, 92 (1986) 306
36. L. Mølhave, G. D. Nielsen, Indoor Air 2 (1992) 65
37. L. Mølhave in *Environmental Toxicants*, 2nd ed, M. Lippmann (ed), Wiley-Interscience, New York, 2000
38. L. Mølhave in *Indoor Air Quality Handbook*, J. Spengler, J. M. Sanmet, J. F. McCarthy (eds), McGraw-Hill, New York, 2001
39. L. Mølhave, Indoor Air 13 (2003) 12
40. K. Anderson, J. V. Bakke, O. Bjørseth, C.-G. Bornehag, G. Clausen, J. K. Hongslo, M. Kjellman, S. Kjærgaard, F. Levy, L. Mølhave,S. Skerfving, J. Sundell, Indoor Air 7 (1997) 78
41. B. Berglund, I. Johansson in *Achieves of the Center of Sensory Research*, B. Berglund, T. Lindvall (eds), Stockholm University and Karolinska Institute, Stockholm. Vol. 3, 1996
42. P. Wolkoff, G. D. Nielsen, Atmos. Environ. 35 (2001) 4404
43. K. Sexton, R. S. Dyer in *Indoor Air and Human Health*, 2nd ed, R. B. Gammge, B. A. Berven (eds), CRC Press Inc., Florida, USA, 1996
44. J. Q. Koenig, *Health Effects of Ambient Air Pollution*, Kluwer Academic Publishers, Maschusetts, USA, 2000
45. E. Shorter, Scand. J. Work Environ. Health 23 (1997) 35
46. C. M. Ryan, L. A. Morrow, J. Consul. Clinic. Psych. 60 (1992) 220
47. T. Godish, *Sick Buildings: Definition, Diagnosis and Mitigation*, Lewis Publishers, Boca Raton, 1995
48. P. S. Burge, Occup. Environ. Med. 61 (2004) 185
49. WHO (World Health Organisation). *Indoor Air Pollutants, Exposure and Health Effects Assessment.* European Reports and Studies No. 78. Copenhagen, Denmark, 1982

50. F. Kondo, Y. Ikai, T. Goto, Y. Ito, H. Oka, H. Nakazawa, Y. Odajima, M. Kamijima, E. Shibata, S. Torii, Y. Miyazaki, Bull. Environ. Contam. Toxicol. 77 (2006) 331
51. J. Bourbeau, C. Brisson, S. Allaire, Occup. Environ. Med. 54 (1997) 49
52. L. A. Wallace in *Indoor AirPollution and Health*, Marcel Dekker, New York, 1997
53. J. J. K. Jaakkola, P. Tuomaala, O. Am. J. Public. Health 84 (1994) 422
54. L. Rothman, M. I. Weintraub, Neurol. Clin. 13 (1995) 405
55. L. Mølhave, Indoor Air 4 (1991) 357
56. L. Mølhave, G. Clausen, B. Berglund, J. De Ceaurriz, A. Kettrup, T. Lindvall, M. Maron, A. C. Pickering, U. Risse, H. Rothweiler, B. Seifert, M. Younes, Indoor Air 7 (1997) 225
57. S. C. Sofuoglu, D. J. Moschandreas, Indoor Air 13 (2003) 332
58. US EPA (United States Environmental Protection Agency) *Introduction to IAQ*, report no. EPA/400/3-91/003 (1991), US Environmental Protection Agency, Washington, D.C.
59. B. T. Burton in *Indoor Air Pollution and Health*, E. J. Bardana, A. Montanaro (eds), Marcel Dekker, New York, 1997
60. E. E. Sandmeyer in *Patty's Industrial Hygiene and Toxicology*, Vol. 2, 3rd ed, G. D. Clayton, F. E. Clayton (eds), Wiley, New York, 1982
61. L. Mølhave, J. Exposure Anal. Environ. Epidemiol. 1 (1991) 63
62. P. Wolkoff, P. A. Clausen, B. Jensen, G. D. Nielsen, C. K. Wilkins, Indoor Air 7 (1997) 92

Chapter 2 Study of Odour and Nasal Pungency

2.0. INTRODUCTION

Chemosensory detection of VOCs by humans rests on the senses of smell and sensory irritation. Several terms have been employed that subsume the sensations evoked by chemicals that are typically viewed as irritative. For example, in 1912, Parker introduced the concept of the "common chemical sense" (CCS) to describe general mucosal sensitivity to chemicals [1,2]. More recently, the terms "chemesthesis" and "pungency" have been used to describe sensations evoked by chemicals that are not properly odours. Such pungency includes: piquancy, tingling, prickling, irritation, stinging, burning, and freshness, among others.

There are two types of odour threshold, detection and recognition. Detection threshold is defined as the lowest concentration of an odorant in clean air at which the odour of the mixture is detectable. It is the awareness of the presence of an added substance. Recognition threshold is defined as the lowest concentration of an odorant in clean air at which the odour of the mixture can be recognized. It is the recognition of a characteristic of a specific odour quality. To understand this concept an analogy is made using other senses. The threshold in sound is demonstrated by increasing the volume of a radio until a sound can be heard (detection threshold). The volume is continued to be increased until speech is understood (recognition threshold).

The sense of smell gives rise to the perception of odours, whereas chemesthesis gives rise to the perception of pungency. Whereas odours are detected in the olfactory mucosa that covers the upper back portion of the nasal cavity via the olfactory nerve (cranial nerve I), chemesthetic sensations are mainly detected in all three mucosa: ocular, nasal, and oral, mainly via the trigeminal nerve (cranial nerve V) [3].

The pharmacological and toxicological characterization of the senses of smell and chemesthesis includes the study of the breadth and sensitivity of responses towards the spectrum of chemicals. Recent studies on the molecular biology of smell [4] have provided additional support to the long-held view of the existence of a large number of different odorant receptors, probably in the order of 1,000, of which about a third are functional. Since a given odorant can activate more than one type of receptor, and since

a given type of receptor can be activated by several odorants, the number of possible odorants that can be recognised may reach several million [5,6,7]. Of course, odour detection is another matter altogether, and within the context of odour detection, a systematic strategy to study the breadth of chemical tuning and sensitivity in olfaction and chemesthesis has considerable merit. One strategy to uncover the physicochemical basis for odour detection thresholds and pungency thresholds of VOCs consists of measuring chemosensory thresholds for homologous series of compounds. Over the past fifteen years, Cometto-Muniz and co-workers have carried out a systematic investigation into thresholds for sensory irritation and odour, using panels of human subjects under carefully controlled conditions and homologous series of compounds, including alcohols, ketones, aldehydes, carboxylic acids, aromatic hydrocarbons, terpenes, among other VOCs.

Previous research showed that olfaction is more sensitive than chemesthesis to detect VOCs [8]. One approach to avoid the influence of smell is to probe nasal chemosensory detection in participants lacking olfaction, that is anosmics. Another approach consists in testing participants with an intact olfaction, that is, normosmics, but in terms of nasal localization or lateralization rather than detection. This technique tests the ability of subjects to identify the nostril receiving the VOC when the contralateral nostril simultaneously receives plain air [9]. Previous research determined that such localization is mediated by trigeminal, not olfactory input [10].

There are more than 100,000 industrial chemicals, and even if only a third could be classed as VOCs or semi-volatile organic compounds, it is obvious that experimental determination of potency towards humans cannot possibly be extended to more than a very small proportion. The use of animal experiments allow the study of VOCs that are too toxic to be tested on humans, but does not help very much as regards the sheer number of compounds examined.

There is therefore a very definite need for some type of prediction of the potency of VOCs towards humans. Even if restricted to VOCs that act through 'physical' mechanisms, rather than through 'chemical' or 'reactive' mechanisms, such predictions would considerably help to fill the gap between the relatively small number of VOCs studied to date, and the very large number of chemicals that could be encountered.

First a summary of the biochemical study of olfaction is presented. Biological information on nasal irritation is also reported. Some methods available to measure odour and sensory irritation are listed. Attention is mainly focused on standardised

methods developed by Cometto-Muniz and Cain to measure odour detection thresholds, nasal pungency thresholds and eye irritation thresholds.

2.1. SENSORY CHANNELS

2.1.1 Olfaction

The nose is a complex structure that provides access of ambient air to the lower respiratory tract. The nose contains the sense of smell, where odorant molecules come in contact with olfactory receptors. The initial events in olfaction consist of the movement of odorants molecules from the environment to the olfactory receptors. In the vertebrate nose, these events include the transport of the molecules of volatile compound to the airspace above the olfactory receptors and then the transfer of at least a portion of these molecules into the olfactory mucosa, sited in the roof of the nasal cavity [11], see Figure 2.1. Once in the olfactory mucosa, these odorant molecules are in position to stimulate the olfactory receptors, which are located on the ciliated olfactory epithelium buried in an aqueous mucus layer, as will be explained in detail in the next section.

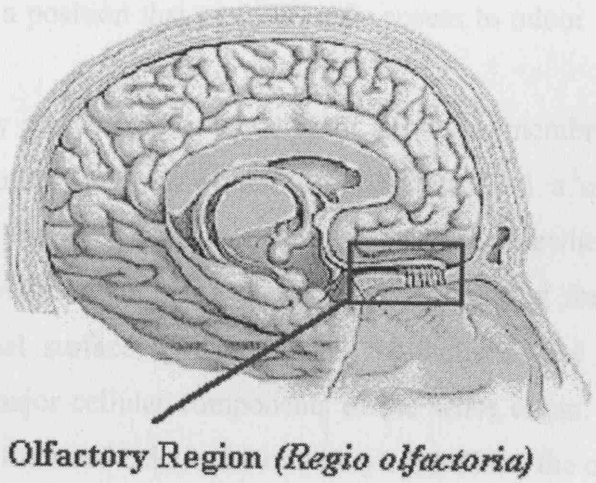


Figure 2.1. Diagrammatic representation of where the olfactory region is located

The sense of smell is based on the ability of the olfactory sensory neurons (OSNs) to respond to environmental chemicals by initiating a nerve impulse. These OSNs have two primary properties: (1) the ability to react to one or perhaps a few specific chemical

molecules, this means that the sensory cell can discriminate among molecules to which it is reactive and others to which it is not, more over it can respond to the intensity of the stimulus as well as its quality; (2) the ability to transform (transduce) the chemical signal into a meaningful response in the olfactory sensory cell, the meaningful response is a nerve impulse that is transmitted to the brain.

The selectivity response to some chemicals and not to others is based on the presence of the receptors to which the odorant chemicals are specifically bound. These receptors are on the olfactory cell surface, where they are accessible to environmental chemicals, linked to a transducing mechanism that ultimately elicits the effect. Thus when an odorant chemical is bound to an appropriate receptor on an olfactory sensory neuron, the transducing system is activated and a nerve impulse is generated. Absence of a receptor for a particular chemical stimulus will preclude the cell from responding to the chemical.

2.1.1.1. Structure of the olfactory system

The olfactory system is incorporated into the respiratory system. The nasal cavity forms the most anterior end of the respiratory system. It is divided into three sectors: a vestibule or entry region, a respiratory region, and an olfactory region. The olfactory region is located in a position that permits ready access to odour molecules reaching it [12].

The olfactory sense organ is made up of a mucous membrane consisting of two major layers: an epithelium, the olfactory epithelium, and a connective tissue, the lamina propia, see Figure 2.2. Like mucous membrane elsewhere in the body, it is maintained in a moist condition by glandular secretions that form a layer of mucus covering the external surface of the olfactory epithelium. The olfactory epithelium contains the three major cellular components of the sense organ: the sensory cell, the supporting cell, and the basal cell. In the olfactory epithelium, the olfactory receptor cell body has a peripheral location in direct contact with the external environment, see Figure 2.2 [12]. A fourth cell type, the microvillar cell, has been seen in the upper region of the olfactory neuroepithelium in humans [13,14].

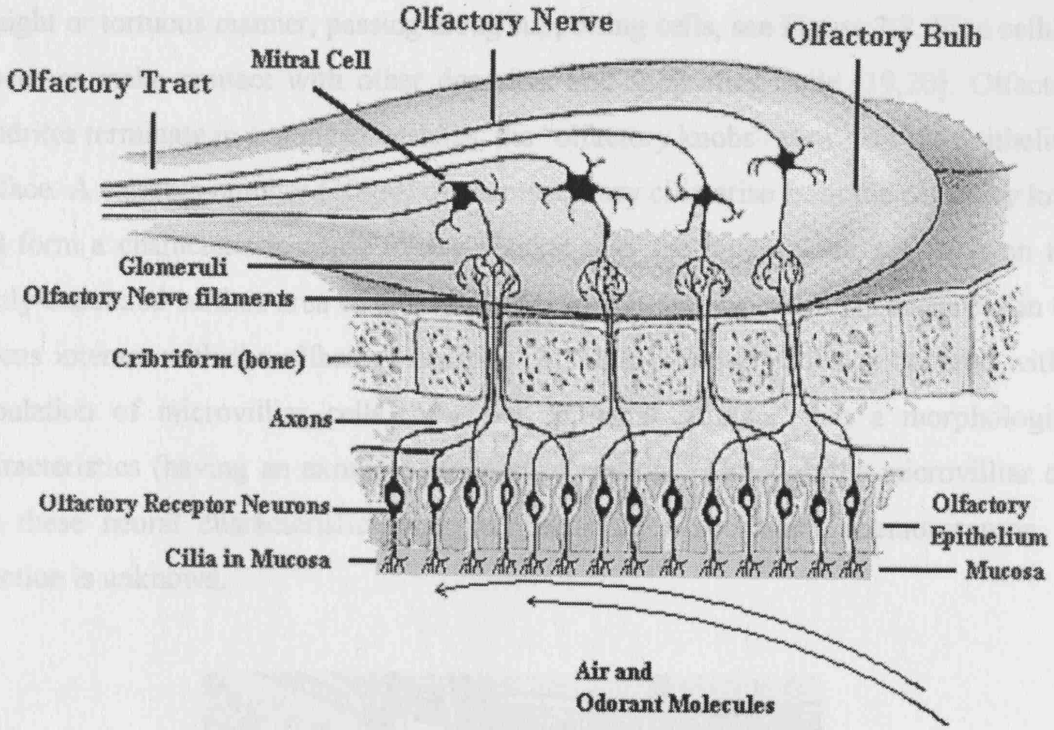


Figure 2.2. Diagrammatic representation of the olfactory system

a) Olfactory receptor neurons

The olfactory receptor neuron is a slender bipolar neuron generally located in the basal two-thirds of the epithelium, see Figure 2.2. Olfactory receptor neurons are in direct contact with others and with adjacent supporting cells [15,16]. They arise from neuroblast cells (basal cells) at the base of the epithelium, see Figure 2.3.

Olfactory neurons develop an axon process that gradually tapers to a 0.1-0.3 μm process from a conical region at the base cell body [17]. Olfactory axons are always unbranched, unmyelinated, and grouped within the epithelium in small fascicles that pass through the basal lamina into the underlying lamina propria. Within the lamina propria, olfactory axons are ensheathed by sheath cells. These cells enclose olfactory axons in an unusual manner in which several mesaxon processes may surround up to 50 axons [15,18]. Olfactory axons then form larger fascicles that pass through the foramina of the ethmoid bone's cribriform plate and enter the olfactory bulb, see Figure 2.2. Olfactory axons form their characteristics terminal structure, the "glomerulus", in the olfactory bulb, where they make their first synapse with dentrites of second-order neurons.

The olfactory dendrite, with a diameter of 1-2 mm, originates from the apical part of the cell body. Dendrites are unbunched and extend to the epithelial surface in a

straight or tortuous manner, passing along supporting cells, see Figure 2.3. Fine cellular processes make contact with other dendrites and supporting cells [19,20]. Olfactory dendrites terminate in expanded vesicles, the "olfactory knobs" at or near the epithelium surface. A variable number (10-60) of chemosensory cilia arise from the olfactory knob and form a characteristic dense ciliary blanket over the mucosal surface. It is on this vastly expanded surface area of ciliary plasma membrane that odorants dissolved in the mucus interact with the olfactory neurons [21,22,23]. Sensory cilia is covered with a population of microvilliar cells, near the epithelial surface, with a morphological characteristics (having an axonlike process) of neurons. Although the microvilliar cell has these neural characteristics and may serve as a secondary chemoreceptor, its function is unknown.

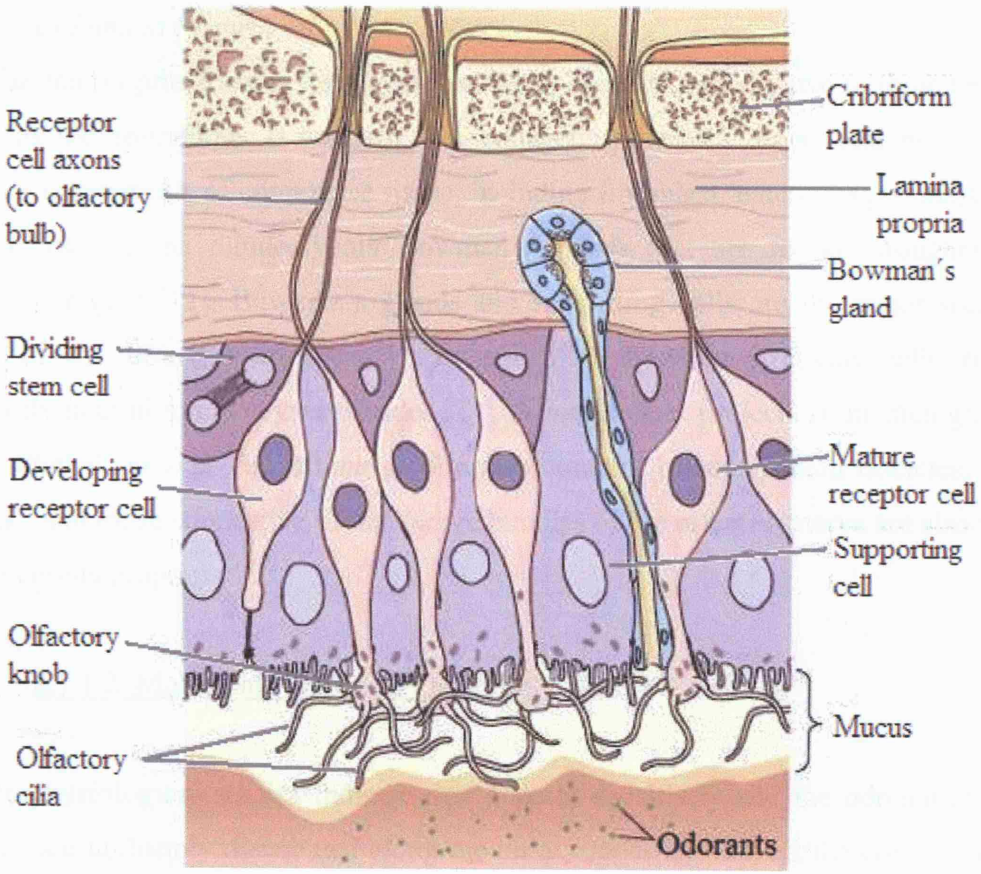


Figure 2.3. Structure of the olfactory epithelium.

b) Supporting Cells

The columnar supporting cells are major secretory components of the olfactory mucosa. They have apical microvillar surfaces, taper basally, and attached at the basal lamina by

footlike process. Their functions may include: (a) insulation of olfactory receptor cells, (b) a glialike function within the neuroepithelium, (c) secretion (of acidic, sulphated, and/or neutral mucopolysaccharides from the apical pole of the cell surface), (d) transepithelial transport of molecules, and (e) guides for developing neurons and their neural process [15,16,19,24,25,26].

c) Basal Cells

Basal cells are irregular shaped and located in the lower epithelium, near the lamina propia. These cells are the stem cells that mitotically divide, given rise to new cells that develop into olfactory receptors neurons [27]. The olfactory system replaces neural elements normally and when injured.

d) Lamina Propria

The lamina propria, located directly below the epithelium, may be two to three times as thick as the epithelium. It contains thin collagen and elastic fibers and most of the cellular components of connective tissue, including fibroblast, macrophages, mast cells, leukocytes and the tubuloalveola Bowman's glands that are found throughout the olfactory region [11]. Bowman's glands and supporting cells are the major secretory units of the olfactory mucosa. Histochemically, Bowman's glands cells contain primarily neutral mucopolysaccharides [21]. A single duct projects from each gland to the epithelial surface. The lamina propria also contains olfactory axon fascicles, blood vessels, and loose connective tissue. Nerve bundles of the olfactory nerve are also found in the lamina propria.

2.1.1.2. Molecular Aspects of Olfaction

Electrophysiological studies indicate that odorant sensitivity and the odorant-induced current are uniformly distributed along the cilia, suggesting that all the components of the immediate responses to odorants are localized to the cilia. ORNs comprise 75-80% of the cells in the epithelium [12]. They are functionally homogeneous, that is they all detect odorants.

The peripheral olfactory system is well adapted structurally to perform its function. As mentioned, the olfactory primary sensory neurons are located in a portion of the olfactory epithelium facilitating their direct contact with inhaled odorants. The

initial events of the olfactory transduction occur in the ciliary plasma membrane over the mucosal surface [28,29], when odorants interact with specific receptors, odorant receptors (ORs), on cilia of ORNs [30,31], see Figure 2.4. In the detection of an odour molecule, the molecule is somehow recognized by the OR, and the process of this chemical recognition leads to the induction of electrical currents in the cell. An initial generator current that is ligand induced is assumed to exits, and it seems that this generator current in turn induces a voltage activated current [32]. In humans the overall number of ORs may be in excess of 1,000, but only one third of them appear to be functional.

In order to reach ORs placed in the olfactory cilia, hydrophobic odorants must pass through an aqueous medium. Odorant binding proteins (OBP) are thought to solubilize the lipophilic molecules, and allow their transport through the aqueous mucus barrier toward the cilia of ORNs. Alternatively, OBP may function to remove odorants from sensory epithelium and cilia.

ORs subsequently couple to an olfactory G-protein (G_{olf}) to activate a second messenger system, adenylyl cyclase (AC) [33,34], see Figure 2.4. Activation of second messenger pathways, such as AC, via an intermediary regulatory component allows additional control over the transduction process and provides amplification, since one liganded receptor can activate multiple G proteins, each of which in turn can cause the generation of many intracellular second messengers molecules.

Cyclic adenosine 3'5'-monophosphate (cAMP) is generated by adenylyl cyclases from adenosine tri-phosphate (ATP). Electrophysiological and biochemical studies confirm that cAMP is the key messenger in the initial phase of odorant detection [35,36,37,38]. An odorant increases cAMP concentration and open an olfactory cyclic nucleotide-gated channel (oCNC), see Figure 2.4, resulting in an influx of Na^+ and Ca^{2+} [39]. The immediate response is the generation of a graded receptor potencial [40]. The increase in Ca^{2+} may also regulate downstream events [41,42]. The ability of CNCs to conduct Ca^{2+} determines both the rise time and the amplitude of the olfactory receptor current, as well as its termination after the stimulus. Odor-induced Ca^{2+} influx through CNCs triggers a Ca^{2+}-gated Cl^- channel and causes a depolarizing Cl^- efflux that amplifies the receptor current [43]. The extrusion of Ca^{2+} ions is mediated by Na^+/Ca^{2+} exchange mechanisms in cilia and probably involves Ca^{2+}-ATPases in knobs, dentrites, and cilia [44,45,46].

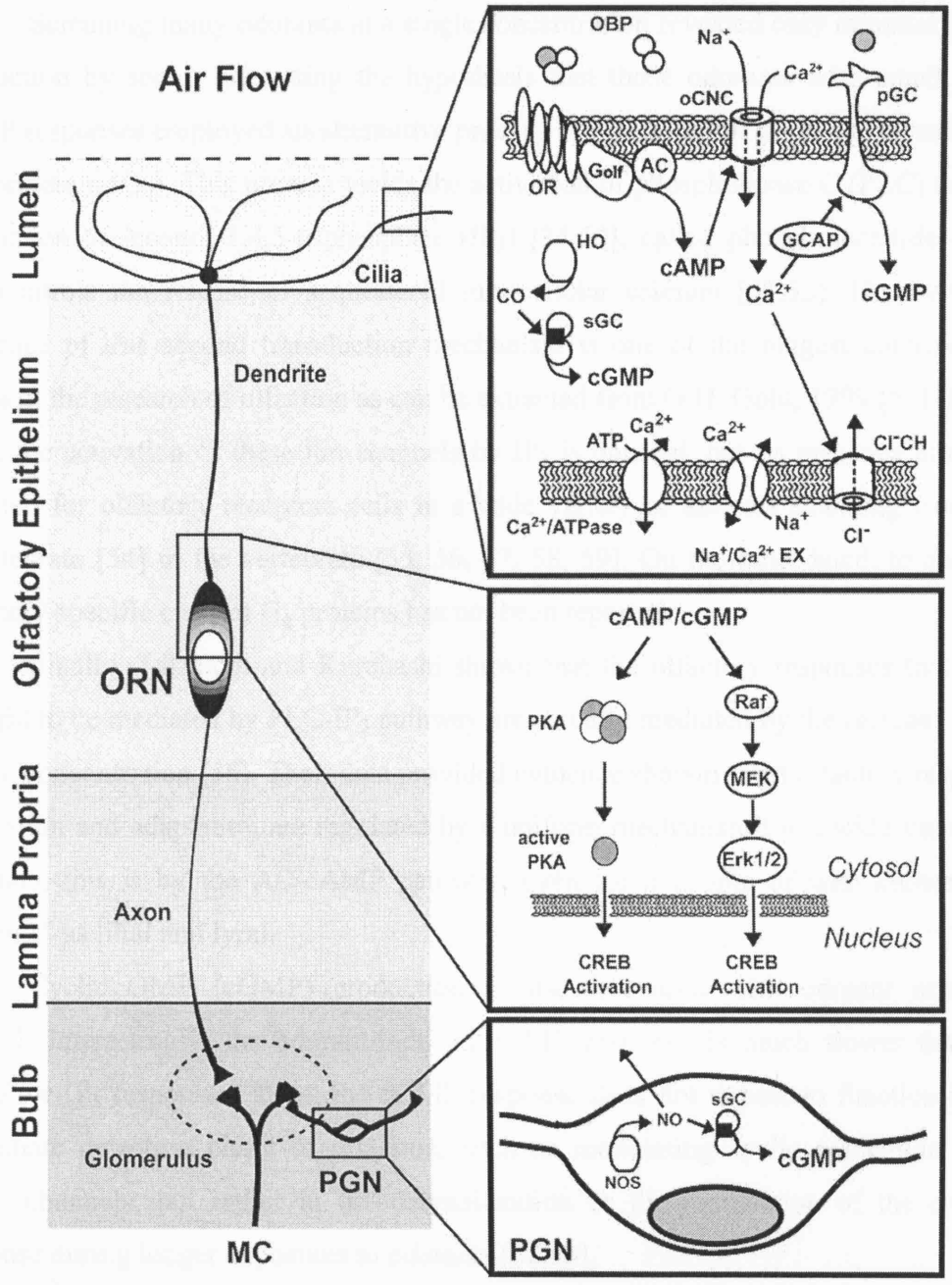


Figure 2.4. Model of odorant signal transduction. Source: G. V. Ronnett, C. Moon, Annual Rev. Physiol. 64 (2002) 189

The activity of oCNCs is regulated in various ways. The most rapid is the negative feedback inhibition of oCNCs by Ca^{2+}/calmodilin. Ca^{2+}/calmodulin causes decrease in the apparent cyclic nucleotide affinity of olfactory CNCs [47]. On the other hand, the increase in Ca^{2+} levels acts to stimulate phosphodiestarase (PDE) activity, and inhibit adenylyl cyclase by phosphorylation, both processes leading to a reduced cAMP concentration [48,49].

Screening many odorants at a single concentration revealed only minimal cAMP production by some, generating the hypothesis that those odorants with small or no cAMP responses employed an alternative process in which an OR couple presumably to a G-protein q (G_q). This process yields the activation of phospholipase C (PLC) and the production of inositol-1,4,5-triphosphate (IP_3) [34,50], called phosphoinositide cycle, that controls the release of sequestered intracellular calcium [51,52]. However, the existence of this second transduction mechanism is one of the biggest controversial issues in the research of olfaction as can be extracted from G.H. Gold, 1999 [53]. In one hand, the activation of these ion channels by IP_3 is unusual, but its presence has been reported for olfactory receptors cells in a wide variety of animals spanning from the invertebrate [54] to the vertebrate [55, 56, 57, 58, 59]. On the other hand, to date, an olfactory specific class of G_q proteins has not been reported.

Finally, Takeuchi and Kurahashi shown that the olfactory responses that were thought to be mediated by PLC-IP_3 pathway are actually mediated by the increase in the cAMP concentration [38]. Their data provided evidence showing that olfactory response generation and adaptation are regulated by a uniform mechanism for a wide variety of odorants, this is by the AC-cAMP pathway, even for a couple of well known "IP_3 odorants" as lilial and lyral.

Cyclic GMP (cGMP) production is also increased with odorant exposure [60,61]. Interestingly, the odorant-induced cGMP response is much slower than the cAMP or IP_3 responses. Thus, the cGMP response does not appear to function in the immediate detection phase of olfaction, such as modulating cyclic nucleotide gated cation channels, but rather in the desensitization or the modulation of the cellular response during longer exposures to odorants [62,63].

2.1.1.3. Theories of olfaction

Over the recent years a number of theories have been proposed in an attempt to explain how humans perceived odours. In other words, which is the mode of interaction between odorant molecules and the olfactory receptor neurons. Included are the steric theory [64,65,66], the vibrational theory [67,68,69], the membrane-diffusion theory [70,71], complexation [72], the chromatography analogy [73], the hydrogen-bond dispersion model [74] and the vibrational tunnelling theory [75]. Here are reviewed the most prominent theories.

a) The Steric Theory of Odour

It was in 1946 when Linus Pauling proposed a theory indicating that a specific odour quality is due to the molecular shape and size of the chemical [64]. The idea was taken up by Moncrieff three years later who proposed the idea of a "Steric Theory of Odour" that stated airborne chemical molecules are smelled when they fit into certain complimentary receptor sites on the olfactory receptor neurons [65]. Subsequently, in 1952, Amoore proposed that there might be only seven primary odours (ethereal, camphoraceous, musky, floral, minty, pungent and putrid) determined by the shape and size of molecules, because of their high frequency of occurrence amongst 600 organic compounds [66]. Although in a modified and refined form this theory has survived to this day, it is no longer tenable. As has been repeatedly pointed out, structure-odour relations provide conflicting evidence [76,77,78,79]. Moreover, in a recent review it is concluded that shape-based structure–odour relations are in a ''sorry state'' [80].

b) The Vibrational Theory of Odour

Vibration theory was introduced in the 1930s by Dyson [67]. He suggested that the infrared resonance (IR), which is a measurement of a molecules vibration, might be associated with odour. This idea was later extended by Wright [68] in the mid 1950s as infrared spectrometers became generally available for such spectral measurements, which Wright was able to correlate with certain odorants [69]. They both argued that there was a better correlation between the odour of a molecule and its vibrational spectrum than between odour and shape. In addition to the original theory, Wright proposed that since optical means were out of questions, detection of molecular vibrations might necessarily be mechanical in nature, and therefore only vibrational modes excited at body temperature could be detected. Nevertheless, no biological mechanism to convert molecular vibration into neuronal activation was proposed, and the theory eventually fell from favour.

c) The Vibrational Induced Electron Tunnelling Spectroscope Theory of Odour

In 1996, Turin provided a mechanism for detection of vibrations by suggesting that receptors were a biological embodiment of an Inelastic Electron Tunnelling Spectroscope (IETS) [75]. This is, instead of accepting the mechanical vibrational spectroscopy theory previously proposed, replaces it with a theory that the receptor cells act as a "biological spectroscope". He proposed that when the olfactory receptor neuron

binds an odorant, electron tunneling could occur across the binding site by exciting its vibrational mode. This only happens if the vibrational mode equals the energy gap between filled and empty electron levels. The electron tunneling then activates a G-protein cascade. Receptors are therefore "tuned" to the vibrational frequency of particular odorants, rather like cones are "tuned" to particular wavelengths of light. For some reason Turin's theory was not tested, even when it seems quite plausible, until 2004 when Keller and Vosshall [81] tested the vibration theory's key psychophysical predictions, as formulated by Turin. They concluded that molecular vibrations alone cannot explain the perceived smell of an odorous molecule. More recently, Brookes and co-workers [82] also assess the proposal by Turin. The conclusion was that humans could recognize odour using phonon-assisted tunneling, provided the receptor has certain properties that are well within ranges known from other biomolecular systems.

2.1.1.4. A Combinatorial Process for Odour Recognition. A New Perspective.

In 1991, Axel and Buck published a study that contributed considerably to our understanding of the molecular basis for the olfactory perception process [83]. They cloned genes, and showed that they belong to an extremely large multigene family that exhibits variability in regions thought to be important in ligand binding. The possibility that each member of this large family is capable of interacting with only one or a small number of odorants provided a plausible mechanism to accommodate the diversity of odour perception. This led to a series of studies in the 1990s by these authors and their co-workers that ended with the awarding of the Nobel Prize in Physiology or Medicine to Linda B. Buck and Richard Axel in 2004.

The conclusion of these studies is that odorant receptors belong to a superfamily of receptors, the G-protein coupled receptors (GPCRs), encoded by a multigene family, whose expression is restricted to the olfactory sensory neurons [83]. They showed that an individual sensory neuron expresses only one of the about 350 receptors genes present in humans [84,30,85] that, although randomly distributed in the epithelium, project to spatially invariant glomeruli in the olfactory bulb generating a topographic map [86,87]. In fact, they showed that neurons expressing different receptors project to different glomeruli, and that the position of each individual glomeruli is topographically defined and is similar for all individuals in a species, leading to the identification of an anatomic olfactory sensory map [88].

They have also demonstrated that this anatomic map is indeed functional and that each odour elicits a sparse pattern of glomerular activation that may confer a signature for different odours in the brain [89,90,91,92]. Moreover, they showed that the projection of the neurons that connect the glomeruli to the protocerebrum of a *Drosophila* (a fruit fly), revealed a spatial representation of glomerular activity in higher brain centres [93,94].

Another interesting finding of their investigations is that the receptor is expressed on both dendrites, where it recognizes odorants in the periphery, and the axons, where it governs target selection in the bulb. And also, that the olfactory receptor plays an instructive role in axon targeting as one component of the guidance process, determining the site of projection in the brain [95].

On the other hand, a very important conclusion of the study carried out by Malnic and co-workers in 1999 is that the olfactory system uses a combinatorial receptor coding scheme, see Figure 2.5, to encode odour identity and to discriminate odours [30].

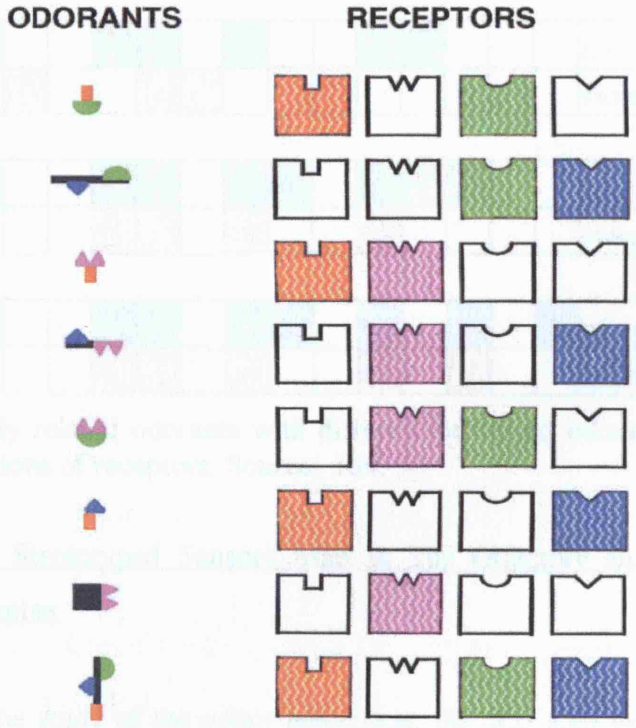


Figure 2.5. Combinatorial Receptor Codes for Odorants. In this model, the receptors shown in colour are those that recognise the odorant on the left. Source: Ref. 30

In this scheme, different odorants are encoded by different combinations of ORs, but each OR may serve as one component of the unique combinatorial receptor codes for

many odorants. Different ORs that recognise the same odorant might recognise different structural features of the odorant.

These studies also demonstrated that changes, sometimes dramatic, in the perceived quality of an odorant that result from an alteration, even a slight one, in its structure (Figure 2.6) or a change in its concentration may be a direct result of changes in its receptor code. At higher concentrations, additional ORs are invariably recruited into the odour response, which may explain why changing the concentration of an odorant can alter its perceived odour. They also suggest in these studies that the size, or complexity, of an odour code might be an important determinant of how easily an odorant can be detected, perhaps reflecting the cumulative intensity of signals transmitted to the olfactory bulb [30].

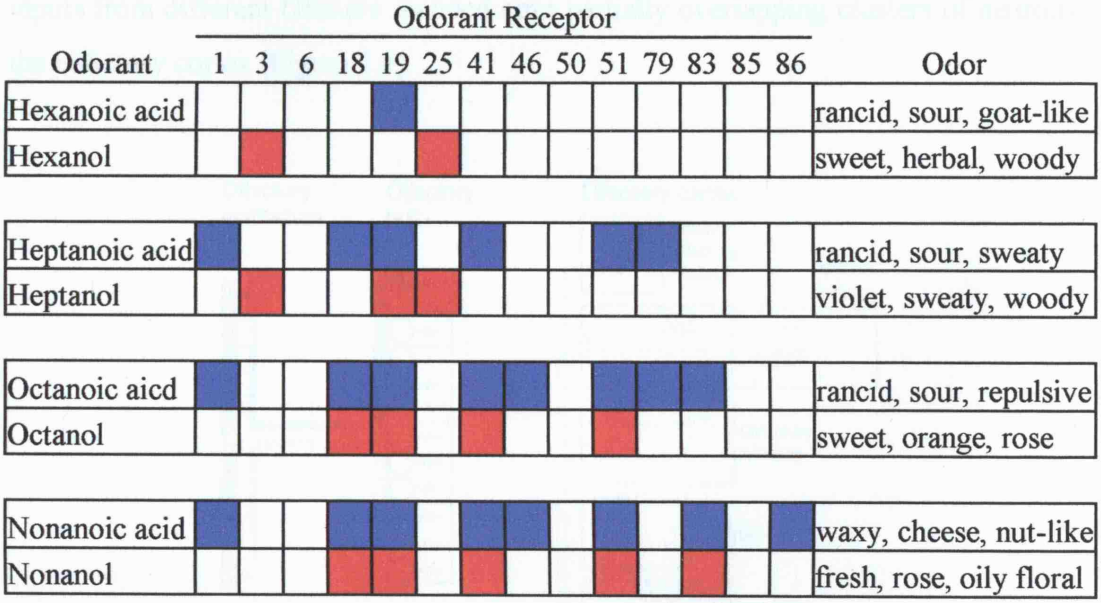


Figure 2.6. Closely related odorants with different perceived odours are detected by different combinations of receptors. Source: Ref. 30

2.1.1.5. A Stereotyped Sensory Map In The Olfactory Bulb And Olfactory Cortex

Continuing with the study of the odour perception, the next step would be to develop the theory of how the brain translates an odorant's combinatorial perception code into a perception.

In 1994, the study by Ressler indicates that the axons of thousands of sensory neurons with the same OR converge in only 2-4 glomeruli, each of which is likely to be

dedicated to one OR [86], in other words, each olfactory sensory neuron in the olfactory epithelium sends a single axon to the olfactory bulb of the brain, see Figure 2.2. Here the sensory axon enters a spherical structure, the glomerulus, where it synapses with the dentrites of bulb neurons, named mitral cells. Each sensory synapses in only one glomerulus. Mitral cells are relay neurons that transmit signals to the olfactory cortex, which is composed of a number of distinct anatomical areas.

There is a stereotyped sensory map in the olfactory cortex [96]. In this map, input from one OR is mapped onto clusters of neurons at a limited number of sites in the olfactory cortex (Figure 2.7). Sites that receive input from a given OR have similar or identical locations among individuals, and most are bilaterally symmetrical in the two hemispheres of the brain. In marked contrast to the olfactory bulb, where inputs from different ORs are segregated in different glomeruli and relay neurons, it seems that inputs from different ORs are mapped onto partially overlapping clusters of neurons in the olfactory cortex (Figure 2.7).

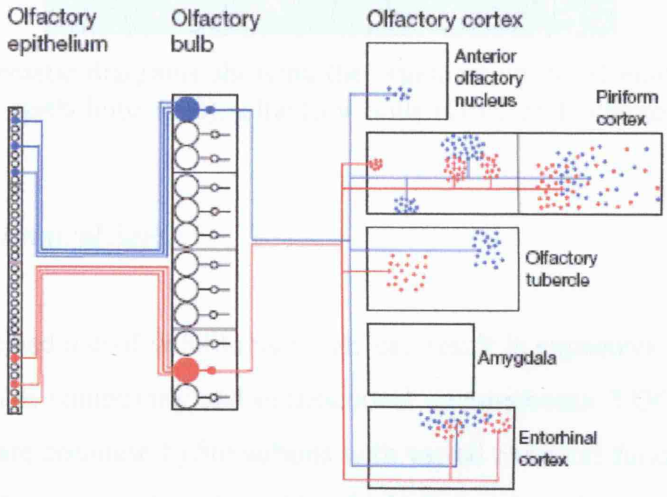


Figure 2.7. Transformations of odorant receptor inputs in the nervous system. Source: Ref. 96

In addition, it is possible that individual cortical neurons receive input from many different ORs as represented in Figure 2.8. Information from the same OR is then processed in parallel in different olfactory cortical areas, potentially allowing OR inputs to be integrated or refined in different ways before delivery to the neocortex and limbic system [96]. Given that different odorants are recognized by different combinations of ORs, such an arrangement could permit an integration of the components of each

odorant's unique receptor code that is important to the generation of diverse odour perceptions.

The existence of a stereotyped map in the cortex suggests a mechanism by which odorants could elicit similar perceptions, and perhaps emotional and physiological responses, in different individuals [96].

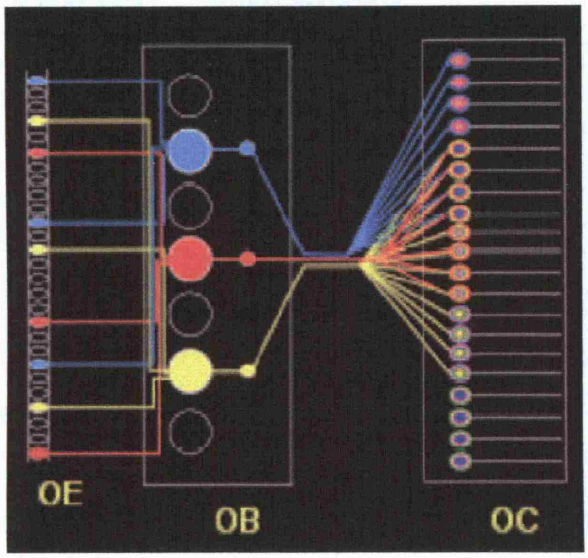


Figure 2.8. Schematic diagrams showing the organization of odorant receptor inputs in the olfactory epithelium (OE), olfactory bulb (OB), and olfactory cortex (OC). Source: Ref. 96

2.1.2 Chemical Sense

The manufacture and use of volatile materials can result in exposures sufficient to cause sensory irritation in community and occupational environments. VOCs found in indoor and outdoor air are common hydrocarbons with varied chemical functionalities such as alcohols, esters, ketones, carboxylic acids, aldehydes, and the like, including linear and branched, saturated and unsaturated, aliphatic and aromatic molecules. Most of them, at high enough concentration, can trigger trigeminal sensory irritation. The word "irritation" can have different meanings. In the present work this word always refers to chemical sensory irritation produced when airborne chemicals stimulate unspecialized free nerve endings. This is distinct from the chemical stimulation of the specialized olfactory receptor cells, see Table 2.1.

In 1912, Parker suggested that free nerve endings in the epithelium are responsible for the detection of chemical irritants and that similar free nerve endings are

located over the entire mucosae [1]. These free nerve endings from exposed upper respiratory tract mucosae, are very susceptible to being stimulated by chemicals, including irritants, given that these sites show high accessibility and permeability.

Table 2.1
Characteristics of the Term "Chemical Sensory Irritation"

Stimuli	Receptive structures	Sensory correlates (examples of sensations)		Other physiological
Airborne chemicals	Free nerve endings, particularly those on readily exposed mucosae: ocular, nasal, oral, and upper respiratory tract	Stinging Piquancy Burning Tingling Freshness Prickling Irritation Itching Cooling	Most of these sensations, if unwanted within a certain context, can be labeled irritative	Nasal airflow changes Breathing pattern changes Changes in secretions Changes in ciliary beat frequency Changes in blood flow Inflammation Release of mediators Electrophysiological responses

Source: Ref. 103

As mentioned above, irritant VOCs comprise a very diverse group of compounds and the structural basis of their activity is not well understood. Nerve recordings [97] and human psychophysical studies [98,99] indicate that response thresholds are strongly related to hydrophobicity within homologous series. The high correlation between hydrophobicity and pungency suggests that irritants must interact with a hydrophobic biophase [100], either the epithelium through which irritants must diffuse or the nerve ending itself.

The term nasal pungency will be used along this work to refer to the chemical sensory irritation evoked by VOCs. As mentioned in the introduction, nasal pungency is detected in the nasal mucosae via the trigeminal nerve (cranial nerve V or CN V), a somatosensory nerve (Figure 2.9). Among the pungent sensations are those of stinging, piquancy, burning, tingling, freshness, prickling, irritation, and the like. As these sensations grow in intensity, they all can be ultimately defined as irritative. Note that pungency includes sensations beyond those simply viewed as irritative (see Table 2.1).

2.1.2.1. Anatomy And Physiology Of The Trigeminal Nerve

The trigeminal nerve (CN V) is the largest of the cranial nerves, being comprised of three major branches: the ophthalmic nerve (CN V1), which conveys purely sensory

information; the maxillary nerve (CN V2), also a pure sensory nerve; and the mandibular nerve (CN V3), which has both a sensory and a motor component (Figure 2.10) [101]. The major sensory innervation to the nasal cavity is from branches of the maxillary division of the trigeminal, nasopalatine. Other sensory branches are from the ophthalmic division, anterior ethmoidal nerve.

The trigeminal nerve fibers reaching the nasal mucosa ramify repeatedly, terminating in free nerve endings [102]. Trigeminal nerve fibers that respond to irritating compounds, contain the neuropeptides substance P (SP) and calcitonin gene-related peptide (CGRP) and extend close to the nasal epithelia surface [103]. Electron microscopic studies suggest that the vast majority of the CN V free nerve endings terminate within the lamina propria. Nevertheless, a few trigeminal fibers terminate at the line of tight junctions only a few micrometers from the surface [103]. For volatile chemicals to stimulate these nerve endings, they must (1) pass into the nasal cavity, (2) partition into and diffuse through the mucus, and (3) cross the epithelial membranes and/or intercellular tight junctions. Since many trigeminal stimulants are lipid soluble, such access is not difficult.

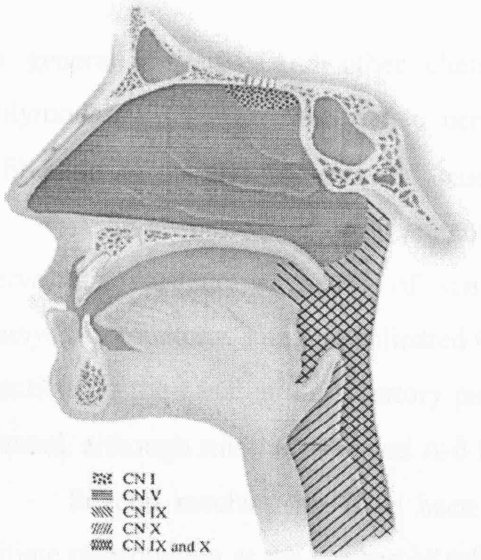


Figure 2.9. Representation of the regions within the nasal and oral cavities innervated by each of several cranial nerves. Source: Ref. 104

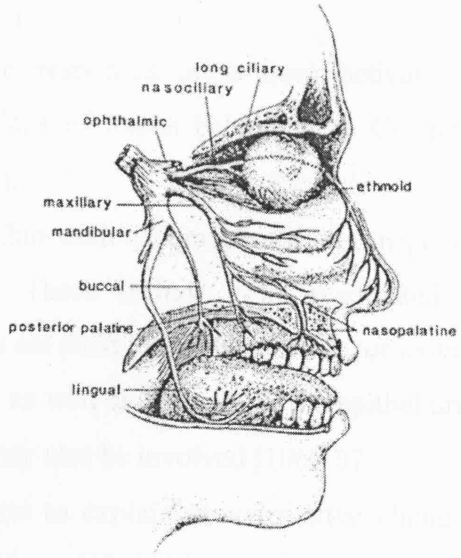


Figure 2.10. Schematic diagram of the branches of the trigeminal nerve that innervate the nasal, oral and ocular epithelia. Source: Ref. 104

The three divisions of the trigeminal nerve come together in an area called the trigeminal ganglion, also termed Gasserion ganglion (Figure 2.11). From there, the trigeminal nerve root continues back towards the side of the brain stem, and inserts into

the pons. Within the brain stem, the signals traveling through the trigeminal nerve reach specialized clusters of neurons called the trigeminal nerve nucleus. Information brought to the brain stem by the trigeminal nerve is then processed before being sent up to the brain and cerebral cortex.

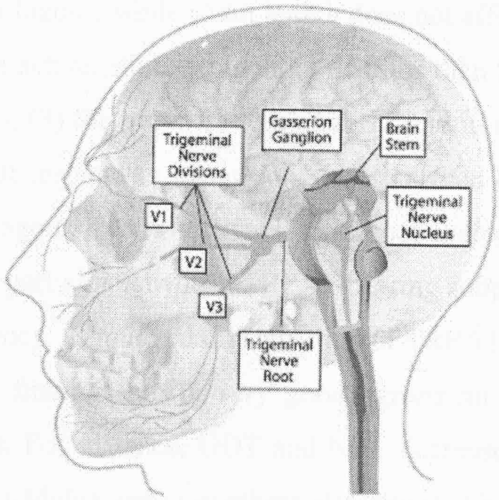


Figure 2.11. Schematic representation of the brain stem.

In general, irritative and other chemesthetic responses arise from activation of polymodal nociceptors within free nerve endings of axons belonging to C- and A-δ fibers, as proposed by Martin and Jessell [105].

A number of fiber types are found within each of branches of the trigeminal nerve that mediate a variety of sensations. These include both myelinated and unmyelinated axons. The unmyelinated C-fibers are most likely responsible for irritative reactions in the nasal and respiratory passages, as well as those from the epithelium in general, although small myelinated A-δ fibers may also be involved [106,107].

Several mechanisms have been proposed to explain how irritative chemicals initiate transduction at the surface of cell membranes [108] although the nature of these processes is still poorly understood. Compounds that are chemically reactive, that is their potency are induced by a chemical mechanism (i.e., covalent bonding with the receptor) [109] can produce irritation directly by reacting with a receptor or indirectly by mucosal tissue damage via chemical reaction without the need to interact with any particular receptor [110]. In the latter case, damaged cells would release endogenous chemicals such as ATP, H^+, and bradykinin (a nonapeptide messenger) which, in turn, could act specifically upon ion channels to produce the neural response [111,112]. On

the other hand, there are other compounds that are likely to act on specific receptors: menthol [113]; capsaicin, the pungent agent in hot papers [106]; and nicotine [114].

Studies of homologous series of irritant VOCs shed light on the contribution of the chemical structure in their interactions with receptors. Inoue and Bryant [115] found that (1) longer chain aldehydes and alcohols activate neurons at lower concentration than their shorter homologous, while chain length does not affect enaldehyde responses; (2) aldehydes are more active, stimulating more neurons than the same concentration of corresponding alcohols; (3) the presence of a double bond in enaldehydes increases the number of neurons that respond to lower concentration than corresponding aldehydes; (4) the three homologous series activate a subpopulation of capsaicin-sensitive nociceptors [116], and partially activate a TRPA1-bearing subpopulation of nociceptors, enough to cause pungency, as mustard oil activation of TRPA1 does [117].

Some of these findings are in very good agreement with those by Cometto-Muñiz and co-workers. For instance, ODT and NPT decreases when the chain length increases [98]. Cometto-Muñiz and co-workers also reported higher threshold values for aldehydes than alcohols [99]. Additionally, these findings are also consistent with those by Alarie and co-workers [109], since they demonstrated that enaldehydes have lower irritation threshold in mice than corresponding saturated aldehydes.

2.2. OVERVIEW OF ODOUR AND NASAL PUNGENCY MEASUREMENT TECHNIQUES

Methods to generate stimuli and deliver them to the nose divide themselves relatively well into the static, where odour vapour exists in an enclosed or semi-enclosed volume, or the dynamic, where vapour flows continuously, carried by dilution gas, e.g. nitrogen or odourless air, toward the nose of a subject [104]. In static systems, an investigator may make the assumption that the number of molecules present in the vapour phase at steady state follows Raoult's law for ideal solutions or Henry's law for ideal dilute solutions. These laws assume proportionality of concentration in the headspace above a solution to molar fraction of solute. A solution 10 times less concentrated than another should have 10-fold fewer molecules in the headspace [118].

Procedures for presenting different concentrations of stimuli to the subject include the draw tube olfactometer of Zwaardemaker; glass sniff bottles; plastic squeeze bottles; air-dilution olfactometers; glass rods, wooden sticks, or strips of blotter paper

dipped in odorant; microencapsulated "scratch and sniff" odorised strips; and bottles from which blasts of saturated air are presented (Figure 2.12) [119].

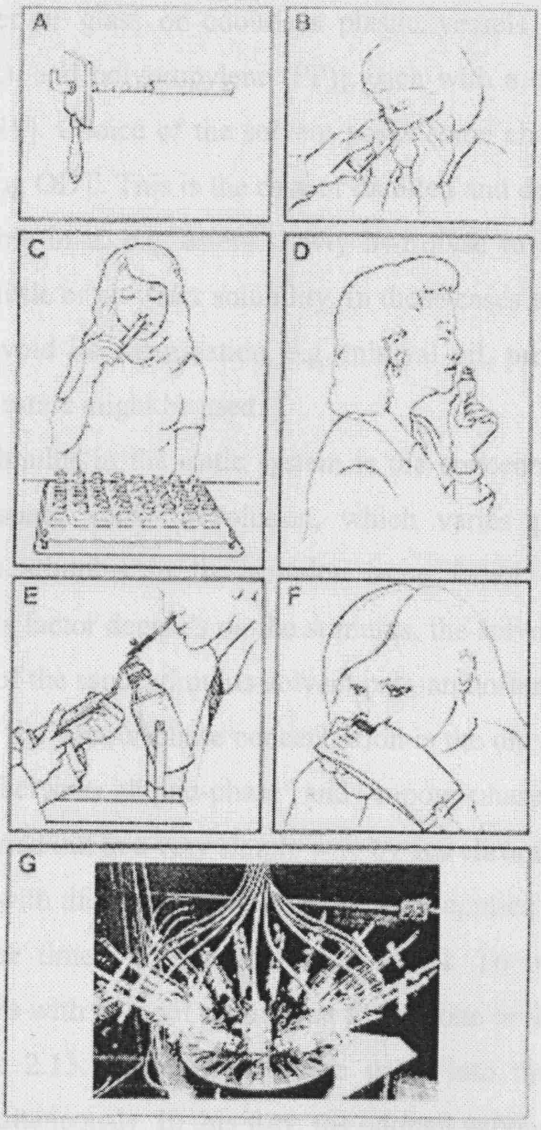


Figure 2.12. Procedures for presenting low concentrations of stimuli to subjects for assessment. A. An early draw-tube olfactometer of Zwaardemaker. B. Sniff Bottle; C. Perfumist's strip; D. Squeeze bottle; E. Blast Injection Device; F. Microencapsulated "scratch and sniff" test; G. Sniff ports on a rotating table connected to the University of Pennsylvania's Dynamic Air-Dilution Olfactometer. Source: Ref. 119

It should be noted that none of these procedures quantify the stimulus at the level of the receptors, as the actual concentration reaching the epithelia can vary as function of factor such as the thickness or composition of the mucus layer, and the shape and size of elements of the nasal chambers (e.g., turbinates), which influence airflow and sorption patterns [104].

2.2.1 Static System

Static olfactometry, generally the easiest form of stimulus control, customarily entails the use of a number of glass or odourless plastic vessels [Teflon®, high density polyethylene (HDPR), and polypropylene (PP)], each with a dilution of odorant in an odourless solvent [119]. Choice of the solvent poses some challenge, as they must be odourless for assessing ODT. This is the case of distilled and deionised water, but some odorants are not stable in it, e.g. esters slowly hydrolyse to acids and alcohols, and many odorants have little or no water solubility. In those cases a solvent with a very low vapour pressure to avoid its vaporization, e.g. mineral oil, propylene glycol, or some low-vapour-pressure esters might be used.

The actual stimulus in the static system is the concentration of the odorant or irritant in the headspace above a solution, which varies proportionally with the concentration of the stimulus in the solution by a factor known as the activity coefficient [120]. This factor depends on the stimulus, the solvent, and sometimes even on the concentration of the same stimulus-solvent pair, and often deviates from Raoult's Law [121]. Checking the vapour-phase concentration is the only way to avoid mistakes about the relation between liquid-phase and vapour-phase concentrations. Such calibration can be carried out in a very simple way by gas chromatography.

One problem with this static system when is not applied correctly is the dilution of the stimulus at the time of sniffing from a vessel. To overcome this problem, squeezable sniff bottles with pop-out spouts that fit into one or the other nostril are used [122,123], see Figure 2.13. Subjects place the spout into the specific nostril, then squeeze and sniff simultaneously. In this way, the odorant vapour enters the nasal cavity efficiently, and each nostril can be tested separately. The simplicity of the method, the ease of use and handling of the bottles, and the possibility of unilateral testing make this procedure convenient in both clinical work [123] and basic research [8,98,99].

Another problem is the time of re-equilibration of the stimulus concentration in the headspace after squeezing the bottle. This problem can be solved simply using a couple of bottles for each concentration step and making an alternative use of each. Additionally, because sniff volumes can be several litres, sniffs that outpace the restoration of saturated stimulus to the vessel become diluted with surrounding air. Thus, the larger the vessels the greater is the sensitivity achieved [124].

Figure 2.13. Squeeze bottles with pop-out spouts

2.2.2 Dynamic System

Dynamic dilution techniques are generally believed to provide a more accurate stimulus concentration than static procedures [104], although they require more complex equipment.

As mentioned above, a dynamic system involves an airstream containing the odorant or irritant substance, generally at vapour saturation for that temperature, that can be mixed in various proportions with odourless air or nitrogen acting as solvent or carrier. An array of tubing, valves, flowmeters, either rotameters or capillaries with nanometers, saturating and mixing vessels, deodorizing and air conditioning (temperature and humidity) devices provide the necessary equipment for the generation and control of odorants [120].

Loading the carrier gas with vapours from liquid stimuli can be achieved by direct vaporization into the carrier stream, by bubbling the odourless gas through the liquid, or by passing the gas over an extended surface of the stimulus (Figure 2.14) [120]. The bubbling method has the disadvantage of forming aerosols, small droplets of the liquid suspended in the gas phase. These can be reduced by aerosol filters (typically glass wool) but may not always be completely eliminated.

2.2.3 Environmental Chambers

Air-dilution olfactometers and static dilution devices may have limitations on the amount of air they deliver to the nose per unit time. These limitations have practical

consequences [124], as said above. However, an odour chamber offers the greatest freedom for human odour and irritation sampling. Such chambers provide a way to allow odour and pungency research with environmental realism [104].

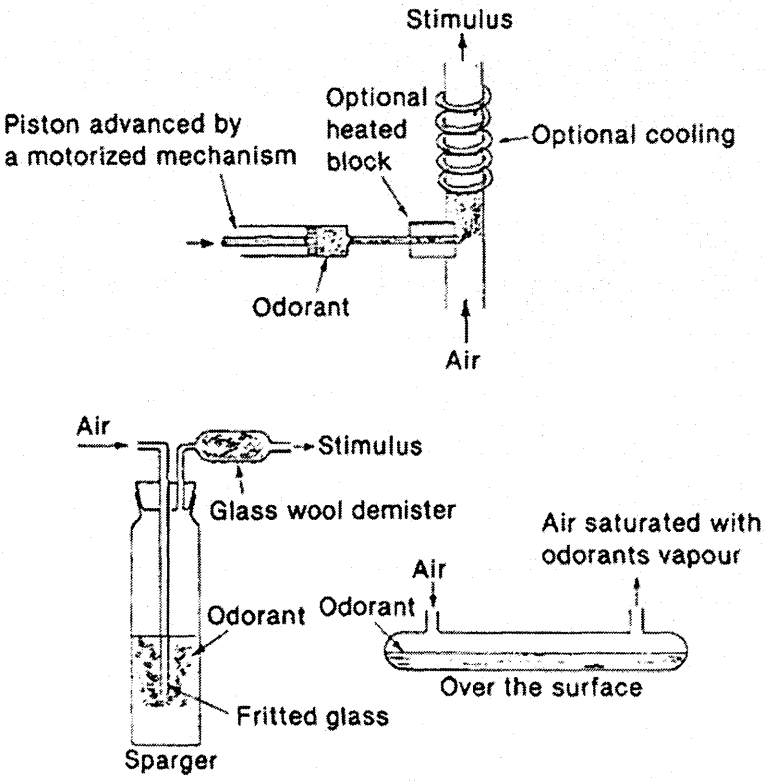


Figure 2.14. Various techniques for introducing odorant into a stream of non-odorous air. Source: Ref. 120

Currently, Cometto-Muñiz and co-workers have focused their work on the use of such chambers. They measure ODT and NPT in this environmental chamber for a wide variety of VOCs, from various homologous series, that have direct relevance to environmentally realistic situation such as indoor air exposures. The aim is to compare these data with those obtained in the last few years by using squeeze bottles and glass vessels. This will allow them to see whether there is a systematic deviation of the data between squeeze bottles and glass vessels, and the environmental chambers. If so, maybe a slight mathematical adjustment could be done to the set of data obtained by the squeeze bottle or glass vessel methods to make them compatible with the data from the chambers, and then continuing working with these experiments that are much more simple and quicker than the environmental chambers.

2.2.4 Methodology

Today, psycophysical methods are commonly employed to assess odour perception and nasal pungency in humans in academic, clinical and industrial settings. Psycophysical procedures can be divided into two categories [104]: threshold procedures (where the goal is to detect barely discernibly stimuli), which will be discussed below, and suprathreshold procedures (where the employed stimuli are clearly discernibly and such indices as perceived intensity are assessed), for further details on this procedure the reader is referred to Doty *et al.* [104].

The threshold measured most often is the detection threshold, also termed absolute threshold, which is the lowest stimulus concentration where the presence of the stimulant is noted. In odour threshold assessment, subjects seek to distinguish the presence of odour from odourless air, and do not seek to specify or recognize odour quality. On the other hand, in nasal pungency assessment, subjects seek to decide which of the multiple presentations produce the stronger sensation.

Odour detection thresholds have been measured since the middle of the nineteenth century [125]. However, a review of the data by different authors reveals non-uniformity in the results, mainly because olfactory thresholds depend largely on the methodology used to gather them and that compilations suffer very much from methodological inconsistencies. Threshold values are generally relative and dependent upon such factors as the method of stimulus dilution, volume of inhalation, species of molecule, type of psychophysical task, and number of trials presented [126]. The findings by Pierce *et al.* [126] emphasize the need for standardization of procedures for threshold testing. Further, it is important to consider that the compounds analysed by different authors are not always the same and the number of them is normally small, which makes even more difficult a proper analysis of data sets obtained by different authors.

In 1989, Cometto-Muñiz and Cain [8] started to explore members of the same chemical series in a search for insight into the physicochemical basis for odour and pungency potency using the same procedure over a long period of time. These authors have also measured other compounds that are not included in chemical series but, because of their presence in indoor air pollution are of particular interest.

2.2.4.1. Ascending Method of Limits (AML) procedure

As previously said, in 1989 Cometto-Muñiz and co-workers started to measure ODT and NPT along homologous chemical series using a uniform procedure. The approach included vapour-phase measurements via gas chromatography, a simple but practical static-dilution delivery system, and a sensory technique based on a multiple forced-choice procedure that controlled for biases and for differences in response criterion across participants [8].

The "absolute" value of odour detection and nasal pungency threshold measured under these standardized conditions might not directly represent thresholds obtained under actual whole-body exposure, but their consistency and the wide variety of VOCs tested proved to be extremely useful to establish robust quantitative structure-activity relationships (QSARs) between potency for odour [127] and associated physicochemical characteristics.

In the squeeze plastic bottle method used by Cometto-Muñiz and co-workers in their earliest work the threshold is obtained by detecting the difference from odour-free background. This is a concentration ascending method, in which a multiple-alternative forced-choice procedure is employed against blanks with increasing concentrations. Therefore, the thresholds reported are nearly equal to the detection threshold. The aim of using an ascending method is to avoid adaptation, i.e. lose of sensitivity from mere stimulation [123].

The method requires participants to select, on each trial, the stronger of two stimuli, one always a blank (solvent) and the other a dilution step of the tested substance, starting with a concentration clearly below threshold. Selection of the blank led, on the following trial, presentation of the next step, i.e. a higher concentration, paired with a blank. Selection of the stimulus led, on the following trial, to presentation of the same dilution step from a duplicate set, also paired with a blank. The procedure continues until the subject selects the substance over the blank five times in a row. The dilution step where this first occurs is taken as the threshold. The ascending concentration approach to the threshold and the alternate use of each nostril helped to minimize the effects of the commonly found phenomenon of olfactory adaptation [128]. Participants in the odour experiments need to perform in the normosmic (i.e., normal sense of smell) range according to a standardized clinical olfactory test described elsewhere [123]. In turn, participants in the nasal pungency experiments need to

perform in the anosmic (i.e., absent sense of smell) range according to the same test, this is to avoid odor biases in their judgments of nasal pungency.

Recently, Wudarski and Doty [129] have compared odour detection thresholds obtained using plastic squeeze bottle with those obtained using glass sniff bottles. In this case a seven-reversal, single staircase (SS) threshold procedure was employed. In this SS procedure, the concentration of the stimulus is increased following trials in which a subject fails to detect the stimulus and decreased following trials where correct detection occurs. Here, the direction of initial stimulus presentation is again made from weak to strong in an effort to reduce adaptation effects of prior stimulation. The finding suggest that, while threshold values using the two presentation procedures can be roughly compared across studies, accurate comparisons may require a slight mathematical adjustment. If this can be considered a general fact, that is the effect is observed in a large variety of date, this difference between odour threshold obtained in one or another way does not interfere with the previously mentioned QSAR [127], except to alter the constant in the equation.

2.2.4.2. The Triangle Odour Bag Method

The triangle odour bag method, the most popular olfactory sensory test in Japan, was first developed in 1972 by the Tokyo metropolitan government [130]. In 1995, this method was introduced into the Offensive Odour Control Law to be used in those cases in which odour complaints cannot be adequately evaluated by the instrumental measurements (gas chromatography or other instruments) of the specific offensive odour substances.

In the triangle odour bag method, the odour threshold is also obtained by detecting the difference from odour-free background [131]. Nevertheless, this is a concentration descending method. Three polyester gas-bags are used, see Figure 2.15. One bag is the odour bag into which a certain amount of the primary odour is injected and the two other bags are filled with only odour-free air. Panel tests to determine the odour concentration are carried out by means of dilution, that is, the test is started with a concentration that the panel can easily detect, and the dilution ratio is successively diluted approximately 3 times in any step of dilution when the answer of the panelist is correct. It is continued until an incorrect answer occurs. In this way, panels can easily detect the odour threshold concentration.

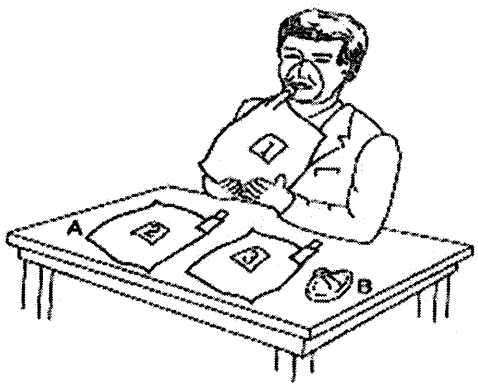


Figure 2.15. Triangle Odour Bag Method Source: Ref. 130

A panel consists of over 6 panellists. It is necessary for each panellist to be over 18 years old. All panellists need also in this procedure to pass a panel screening test described elsewhere [132].

2.3. PSYCOPHYSICAL RESPONSES TO CHEMICAL MIXTURE

In everyday exposures, the human olfactory and the trigeminal chemosensory modalities have to respond to mixtures of VOCs rather than to individual components. However, the vast majority of chemosensory studies deal with single chemicals, and there is very little information on how the sense of smell and chemesthesis process mixture of compounds. Therefore, in 1989, Cometto-Muñiz and co-workers started the study of odour detection and trigeminal chemoreception of binary mixture by measuring of the sensory intensity of the chemicals (threshold) [133].

The findings of these studies reveal that the relation between the total nasal perceived intensity of binary mixture of pungent odorants and that of their components is concentration-dependent [133]. This phenomenon is the result of the simultaneous activation of two sensory channels, olfaction and the common chemical sense, with different rules of additivity for processing mixture of stimuli [134]. This additivity increases with concentration, that is, at low perceived intensities, mixtures are hypoadditive (i.e. lower than the sum of the component intensities); at intermediate intensities, the mixtures are additive; and at high intensities, the mixtures are hyperadditive (i.e. higher than the sum of the component intensities) [133]. A model proposed for the odour and pungency of mixture states that a steeper psychophysical

function leads to a higher degree of additivity in mixtures, with pungency presenting steeper functions and a higher degree of additivity than odour [134,135].

Those works only allowed for a restrictive interpretation of the results. Thus, Cometto-Muñiz and co-workers decided to use a systematic testing of selected VOCs, representative of particular structural and physicochemical properties, by measuring complete concentration-detection (i.e. psychometric or detectability) functions for nasal pungency and odour of the chemicals singly and mixed [136,137]. These functions describe the probability of chemosensory detection as a function of concentration. Information gathered via psychometric functions can be specially revealing in terms of understanding the detection of mixtures. In particular, in terms of the role that detectability level of individual components might play on the degree of response-addition seen in the mixtures [138]. The results from these studies also indicated chemosensory agonism, in the sense of dose additivity, for both olfaction and nasal chemesthesis, but in a different way. At relatively low levels of detectability of the single compounds, the mixtures show complete sensory agonism (i.e. do not deviate systematically from the detectability of each substance presented by itself). In contrast, at relatively high levels of detectability, the mixtures show partial (or incomplete) sensory agonism (i.e. the detectability of the mixture is significantly lower than that of each single substance by itself). This decrease in additivity is more pronounced for olfactory than for trigeminal detection. These results on mixture also lead to the conclusion that the stronger components make a disproportionately large contribution to total chemosensory impact [138].

The outcome also suggested that the analysis of physicochemical descriptors to describe odour and nasal trigeminal detection, as has been used to describe and predict human odour [127] and nasal trigeminal responses [139], will have relevance for mixtures as well [137].

2.4. REFERENCES

1. G. H. Parker, J. Acad. Nat. Sci. Phila. 15 (1912) 221
2. C. A. Keele, Arch. Int. Pharmacodyn. Ther. 139 (1962) 547
3. B. Bryant, W. L. Silver (2000) Chemesthesis: the common chemical senses. In T. E. Finger, W. L. Silver, and D. Restrepo (eds), *The Neurobiology of Taste and Smell*, 2nd ed., Wiley-Liss, New York, p. 73

4. K. J. Ressler, S. L. Sullivan, L. B. Buck, Curr. Opin. Neurobiol. 4 (1994) 588
5. S. Firestein, Curr. Biol. 6 (1996) 666
6. R. R. Reed, Semin. Cell Biol. 5 (1994) 33
7. K. Mori, Y. Yoshihara, Progr. Neurobiol. 45 (1995) 585
8. J. E. Cometto-Muñiz, W. S. Cain, Physiol. Behav. 48 (1990) 719
9. J. E. Cometto-Muñiz, W. S. Cain, Int. Arch. Occup. Environ. Health. 71 (1998) 105
10. G. Kobal, S. Van Toller, T. Hummel, Experientia 45 (1989) 130
11. P. P. C. Graziadei, Tissue Cell 5 (1973) 113
12. Albert I. Farbman, Cell biology of olfaction, NY, 1992
13. D. T. Moran, J. C. Rowley, B. W. Jafek, M. A. Lovell, J. Neurocytol. 11 (1982) 721
14. E. E. Morrison, R. M. Constanzo, J. Comp. Neurol. 297 (1990) 1
15. J. A. Rafols, T. V. Getchell, Anat. Rec. 206 (1983) 87
16. R. M. Constanzo, E. E. Morrison, J. Neurocytol. 18 (1989) 381
17. S. Ramón y Cajal (2000) *Texture of the Nervous System of Man and the Vertebrates*, Vol. 2, English Edition, Springer-Verlag, Wien
18. A. I. Farbman, L. M. Squinto, Dev. Brain Res. 19 (1985) 205
19. P. P. C. Graziadei, Z. Zellforsck. Mikrosh. Anat. 118 (1971) 449
20. W. Breipohl, H. J. Laugwitz, N. Bornfeld, J. Anat. 117 (1974) 89
21. T. V. Getchell, F. L. Margolis, M. L. Getchell, Prog. Neurobiol. 23 (1984) 317
22. D. Lancet, Ann. Rev. Neurosci. 9 (1986) 329
23. R. R. H. Anholt, Am. J. Physiol. 257 (1989) C1043
24. T. V. Getchell, Brain Res. 123 (1977) 275
25. D. Trotier, P. MacLeod, Brain Res. 374 (1986) 205
26. E. E. Morrison, R.M. Constanzo, J. Neurocytol. 18 (1989) 393
27. P. P. C. Graziadei, G.A. Monti Graziadei, J. Neurocytol. 8 (1979) 1
28. T. V. Getchell, Physiol. Rev. 66 (1986) 772
29. G. Lowe, G. H. Gold, Nature 366 (1993) 283
30. B. Malnic, J. Hirono, T. Sato, L. B. Buck, Cell 96 (1999) 713
31. L. B. Buck, Annu. Rev. Neurosci. 19 (1996) 517
32. N. Suzuki (1988) Voltage and Cyclic Nucleotide-Gated Currents in Isolated Olfactory Receptors Cells. In J. G. Brand, B. G. Green, J. R. Mason, M. R. Kare, C. J. Wysocki, M. I. Friedman and M. G. Tordoff (eds.), *Chemical Senses*, Vol. 1, Receptor events and Transduction in Taste and Olfaction, Marcel Dekker, Inc., New York, p. 469

33. U. Pace, E. Hanski, Y. Salomon, D. Lancet, Nature 316 (1985) 255
34. G. V. Ronnett, H. Cho, L. D. Hester, S. R. Wood, S. H. Snyder J. Neurosci. 13 (1993) 1751
35. H. Breer, I. Boekhoff, E. Tareilus, Nature 345 (1990) 65
36. L. L. Brunet, G.H. Gold, J. Ngai, Neuron 17 (1996) 681
37. S.T. Wong, K. Trinh, B. Hacker, G. C. Chan, G. Lowe, A. Gaggar, Z. Xia, G. H. Gold, D. R. Storm, Neuron 27 (2000) 487
38. H. Takeuchi, T. Kurahashi, J. Gen. Physiol 122 (2003) 557
39. S. Firestein, F. Werblin, Science 244 (1989) 79
40. T.V. Getchell, G.M. Shepherd, J. Physiol. 282 (1978) 541
41. S. Frings, R. Seifert, M. Godde, U.B. Kaupp, Neuron 15 (1995) 169
42. U. B. Kaupp., Trends Neurosci. 14 (1991) 150
43. G. Lowe, G. H. Gold, Nature 366 (1993) 283
44. V. E. Dionne, J. Gen. Physiol. 112 (1998) 527
45. J. Noe, E. Tareilus, I. Boekhoff, H. Breer, Neurochem. Int. 30 (1997) 523
46. J. Reisert, H. R. Matthews, J. Gen. Physiol. 112 (1998) 529
47. T. Kurahashi, A. Menini, Nature 385 (1997) 725
48. F. Muller, W. Bonigk, F. Sesti, S. Frings, J. Neurosci. 18 (1998) 164
49. J. Bradley, J. Risert, S. Frings, Curr. Opin. Neurobiol. 15 (2005) 343
50. M. Schandar, K.L. Laugwitz, I. Boekhoff, C. Kroner, T. Gudermann, G. Schultz, H. Breer, J. Biol. Chem. 273 (1998) 16669
51. C. P. Downes, Trends Neurosci. 6 (1983) 313
52. P. W. Majerus, T. M. Connolly, H. Deckmyn, T. S. Ross, T. E. Bross, H. Ishii, V. S. Bansal, D. B. Wilson, Science 234 (1986) 1519
53. G. H. Gold, Annu. Rev. Physiol. 61 (1999) 857
54. D. A. Fadool, B. W. Ache, Neuron. 5 (1992) 907
55. D. Restrepo, T. Miyamoto, B. P. Bryant, J. H. Teeter, Science 249 (1990) 1160
56. D. Schield, F. W. Lischika, D. Restrepo, J. Neurophysiol. 73 (1995) 429
57. N. Suzuki (1994) IP3-activated ion channel activities in olfactory receptor neurons from different vertebrate species. In K. Kurihara, N. Suzuki and H. Ogawa (eds.), *Olfaction and Taste XI*, Springer-Verlag, Tokyo, p. 173
58. M. Kashiwayanagi, K. Shimano, K. Kurihara, Brain Res. 738 (1996) 222
59. R. Kaur, X. O. Zhu, A. J. Moorhouse, P. H. Barry, J. Membr. Biol. 181 (2001) 91
60. T. Ingi, J. Cheng, G.V. Ronnett, Neuron 16 (1996) 835

61. A. Verma, D. J. Hirsch, C. E. Glatt, G. V. Ronnett, S. H. Snyder, Science 259 (1993) 381
62. H. Breer, T. Klemm, I. Boekhoff, NeuroReport 3 (1992) 1030
63. C. Moon, P. Jaberi, A. Otto-Bruc, W. Baehr, K. Palczewski, G. Ronnett, J. Neurosci. 18 (1998) 3195
64. L. Pauling, Chem. Eng. News 24 (1946) 1375
65. R. W. Moncrieff, Am. Perfumer, 54 (1949) 453
66. J. E. Amoore, Perfum. Essent. Oil Rec. 43 (1952) 321
67. G. M. Dyson, Chem. Ind., 57 (1938) 647
68. R. H. Wright, J. Appl. Chem. 4 (1954) 611
69. R. H. Wright, The Sense of Smell, CRC Press, Boca Raton, FL (1982)
70. J. T. Davies, F. H. Taylor, Nature 174 (1954) 693
71. J. T. Davies, F. H. Taylor, Biol. Bull. 117 (1959) 222
72. B. Rosenberg, T. N. Misra, R. Switzer, Nature 217 (1968) 423
73. M. M. Mozell, J. Gen. Physiol. 56 (1970) 46
74. M. Chastrette, D. Zakarya, (1991) in *The Human sense of smell*, D. G. Laing, R. L. Doty, W. Breipohl (eds), Springer-Verlag, p. 77
75. L. Turin, Chem. Senses 21 (1996) 773
76. M. G. J. Beets, (1971) Olfactory response and molecular structure. In L. M. Beidler (ed.), *Handbook of Sensory Physiology*, Vol.4, Springer-Verlag, Berlin, p. 257
77. H. L. Klopping, J. Agric. Food. Chem. 19 (1971) 999
78. G. Ohloff, Experientia 42 (1986) 271
79. P. Weyerstahl, J. Prakt. Chem. Zeitung 336 (1994) 95
80. G. Fráter, J. A. Bajgrowicz, P. Kraft, Tetrahedron 54 (1998) 7633
81. A. Keller, L. B. Vosshall, Nature Neuroscience 7 (2004) 337
82. J. C. Brookes, F. Hartoutsiou, A. P. Horsfield, A. M. Stoneham, Phys. Rev. Lett. 98 (2007) 038101
83. L. Buck, R. Axel, Cell 65 (1991) 175
84. A. Chess, I. Simon, H. Cedar, R. Axel, Cell 78 (1994) 823
85. B. Malnic, P. A. Godfrey, L. B. Buck, Proc. Natl. Acad. Sci. 101 (2004) 2584
86. K. J. Ressler, S. L. Sullivan, L. B. Buck, Cell 79 (1994) 1245
87. R. Vassar, S. K. Chao, R. Sitcheran, J. M. Nunez, L. B. Vosshall, R. Axel, Cell 79 (1994) 981

88. P. Mombaerts, F. Wang, C. Dulac, S. K. Chao, A. Nemes, M. Mendelsohn, J. Edmondson, R. Axel, Cell 87 (1996) 675
89. J. Joerges, A. Juttner, C. G. Galizia, R. Menzel, Nature 387 (1997) 285
90. B. D. Rubin, L. C. Katz, Neuron, 23 (1999) 499
91. N. Uchida, Y. K. Takahashi, M. Tanifuji, K. Mori, Nat. Neurosci. 3 (2000) 1035
92. M. Meister, T. Bonhoeffer, J. Neurosci. 21 (2001) 1351
93. M. Ng, R. D. Roorda, S. Q. Lima, B. V. Zemelman, P. Morcillo, G. Miesenbock, Neuron 36 (2002) 463
94. J. W.Wang, A. M.Wong, J. Flores, L. B. Vosshall, R. Axel, Cell 112 (2003) 271
95. G. Barnea, S. O'Donnell, F. Mancia, X. Sun, A. Nemes, M. Mendelsohn, R. Axel, Science Brevia 304 (2004) 1468
96. Z. Zou, L. F. Horowitz, J. P. Montmayeur, S. Snapper, L. B. Buck, Nature 414 (2001) 173
97. W. L. Silver, J. R. Mason, D. A. Marshall, J. A. Maruniak, Brain Res. 333 (1985) 45
98. J. E. Cometto-Muñiz, W. S. Cain, Indoor Air 4 (1994) 140
99. J. E. Cometto-Muñiz, W. S. Cain, M. H. Abraham, Exp. Brain. Res. 118 (1998) 180
100. M. H. Abraham, R. Kumarsingh, J.E. Cometto-Muñiz, W. S. Cain, Toxicol. In Vitro 12 (1998) 403
101. K. L. Moore, A. F. Dalley (2006) *Clinically Oriented Anatomy*, 5th ed., Lippincott Williams & Wilkins, London
102. N. K. Cauna, K. Hinderer, R. T. Wentges, Am. J. Anat. 124 (1969) 187
103. T. E. Finger, V. St. Jeor, J. C. Kinnamon, W. L. Silver, J. Comp. Neurol. 294 (1990) 293
104. R. L. Doty, J. E. Cometto-Muñiz, A. A. Jalowayski, P. Dalton, M. Kendal-Reed, M. Hodgson, Crit. Rev. Toxicol. 34 (2004) 85
105. J. H. Martin, T. M. Jessell, (1991) Modality coding in the somatic sensory system. In E.R. Kandel, J.H. Schwartz, T.M. Jessell (eds.), *Principles of neural science*, 3rd ed., New York: Elsevier, p. 341
106. N. Jancso, A. Jancso-Gabor, J. Szolcsanyi, Br. J. Pharmacol. Chemother. 31 (1967) 138
107. L. Lundblad, J. M. Lundberg, E. Brodin, A. Anggard, Acta Otolaryngol. 96 (1983) 485
108. G. D. Nielsen, Crit. Rev. Toxicol. 21 (1991) 183

109. Y. Alarie, M. Schaper, G. D. Nielsen, M. H. Abraham, Arch Toxicol 72 (1998) 125
110. G. D. Nielsen, Crit. Rev. Toxicol. 21 (1991) 183
111. P. Cesare, P. McNaughton, Curr. Opin. Neurobiol. 7 (1997) 493
112. E. W. McCleskey, M. S. Gold, Annu. Rev. Physiol. 61 (1999) 835
113. R. Eccles, M. S. Jawad, S. J. Morris, Pharm. Pharmacol. 42 (1990) 652
114. H. Alimohammadi, W. L. Silver, Chem. Senses 25 (2000) 61
115. T. Inoue, B. P. Bryant, Pain 117 (2005) 193
116. M. Schmelz, R. Schmidt, A. Bickel, H. O. Handwerker, H. E. Torebjork, J. Neurosci. 17 (1997) 8003
117. S. E. Jord, D. M. Bautista, H. H. Chuang, D. D. McKemy, P. M. Zygmunt, E. D. Hogestatt, I. D. Meng, D. Julius, Nature 427 (2004) 260
118. J. E. Cometto-Muñiz, W. S. Cain, M. H. Abraham, Chem. Senses 28 (2003) 467
119. R. L. Doty, D. G. Laing (2003) Psycophysical Measurement of Human Olfactory Function, Including Odorant Mixture Assesment. In R. L. Doty (ed.), *Handbook of Olfaction and Gustation*, Marcel Dekker, New York, p. 203
120. W. S. Cain, J. E. Cometo-Muñiz, R. A. de Wijk, (1992) Techniques in the quantitative study of human olfaction. In M. J. Serby and K. L. Chobor (eds.), *Science of Olfaction*, Springer-Verlag, New York, p. 279
121. H. G. Haring, (1974). Vapor pressures and Raoult's Law deviations in relation to odor enhancement and suppression. In A. Turk, J. W. Johnston, Jr., D. G. Moulton (eds.), *Human responses to environmental odors*, Academic Press, New York, p. 199.
122. W. S. Cain, J. Gent, F. A. Catalanotto, R.B. Goodspeed, Am. J. Otolaryngol. 4 (1983) 252
123. W. S. Cain, Ear. Nose Throat J. 68 (1989) 316
124. J. E. Cometto-Muñiz, W. S. Cain, T. Hiraishi, M. H. Abraham, J. M. R. Gola, Chem. Senses 25 (2000) 285
125. W. S. Cain, (1978) History of research on smell. In E. C. Carterette and M. P. Friedman (eds.), *Handbook of Perception*, Vol. 6A, Tasting and Smelling. New York: Academic, p. 197
126. J. D. Pierce, R. L. Doty, J. E. Amoore, Percept. Mot. Skills 82 (1996) 451
127. M. H. Abraham, J. M. R. Gola, J. E. Cometto-Muñiz, W. S. Cain, Chem. Senses 27 (2002) 95

128. W. S. Cain, ASHRAE Transactions 80 (1974) 53.
129. T. J. Wudarski, R. L. Doty, Percept. Mot. Skills 98 (2004) 192
130. Y. Iwasaki (2003) The History of Odor Measurement in Japan and Triangle Odor Bag Method. In *Odor measurement review*. Tokyo: Office of Odor, Noise and Vibration. Environmental Management Bureau, Ministry of Environment, pp. 37-47
131. Y. Nagata (2003) Measurement of odor threshold by triangle odor bag method. In *Odor measurement review*. Tokyo: Office of Odor, Noise and Vibration. Environmental Management Bureau, Ministry of Environment, pp. 118-127
132. K. Saiki (2003) Standard odors for selection of panel members. In *Odor measurement review*. Tokyo: Office of Odor, Noise and Vibration. Environmental Management Bureau, Ministry of Environment, pp. 102-105
133. J. E. Cometto-Muñiz, M. R. García-Medina, A. M. Calviño, Chem. Senses 14 (1989) 163
134. J. E. Cometto-Muñiz, S. M. Hernandez, Percept. Psychophys. 47 (1990) 391
135. J. E. Cometto-Muñiz, W. S. Cain, H. K. Hudnell, Percept. Psychophys. 59 (1997) 665
136. J. E. Cometto-Muñiz, W. S. Cain, M. H. Abraham, J. M. R. Gola, Physiol. Behav. 67 (1999) 269
137. J. E. Cometto-Muñiz, W. S. Cain, M. H. Abraham, J. M. R. Gola, Toxicol. Sci. 63 (2001) 233
138. J. E. Cometto-Muñiz, W. S. Cain, M. H. Abraham, Indoor Air 14 (2004) 108
139. M. H. Abraham, J. M. R. Gola, J. E. Cometto-Miñiz, W. S. Cain, Indoor Built. Environ. 10 (2001) 252

3.0. INTRODUCTION

Irritation is a typical physiological response to a chemical or physical stimulus. Eye irritation is defined as "the magnitude of any stinting, scratching, burning, or other irritating sensation from the eye" [1]. It is one of the most common symptoms reported on indoor air quality studies. Eye irritation is considered to be a transient, reverse and severe symptom. It is related to a destabilization of the outer-eye tear film. It is known that direct installation of compounds may induce a reduction in surface tension, reduce the lipid layer and break-up time, or enhance epithelial damage of conjunctiva.

A key issue in understanding human chemosensory perception involves knowledge of the relevant structural and physicochemical properties of chemicals that govern potency, i.e., absolute detection, and perceptual quality. Within biochemical fields such as pharmacology and toxicology, analogous issues are often addressed as (quantitative) structure-activity relationships (QSARs or SARs). In the taste modality, where there is relative consensus on the number and types of qualitative categories, namely, sweet, bitter, salty, sour, and umami, the approach has involved the search for critical molecular features among families of tastants evoking one or another of these prototypical tastes [2,3,4,5,6]. In the irritancy modality, no such simple and convenient approach can be followed [7]. A practical tactic to overcome this deficiency can be illustrated by considering the study of olfactory responses along members of homologous chemical series, where carbon chain length becomes the "unit of chemical change" that serves as a probe to measure a concomitant olfactory outcome, for example, a detection threshold [8]. A similar tactic can be applied to develop QSARs for detection of chemically induced somesthesis, that is, chemically stimulated feel [9]. Investigators in the chemical senses have labelled this particular kind of chemical sensitivity chemesthesis [10,11], but it is also referred to as chemical nociception or irritation [12,13]. This chemosensory modality plays an important role as a warning system against exposure to deleterious vapours, and, as such, constitutes an early indication of potentially ensuing, more severe, toxicological effects [14]. From this role,

the advantage of being able to predict the chemesthetic potency of airborne chemicals for humans on the basis of chemical structure and properties becomes clear.

3.1. SOURCES AND PATHWAY FOR EYE IRRITATION

Certain types of gaseous indoor pollutants are associated with changes in the tear film structure [15]. Lipophilic and surface-active lipophilic pollutants have been hypothesized to cause thinning of the tear film and destabilized eye tear film and, finally result in reduced break-up time and epithelial damage of the conjunctiva [16]. Chemically non-reactive VOCs (e.g. heptane, toluene, and pentanol) in indoor air are not normally considered to be a direct cause of eye irritation, because their irritation estimates are usually orders of magnitude higher than typical indoor concentrations [17]. However, there are several examples, such as new built houses, painting of buildings, carpet factories, and other more daily situations such as cleaning and cooking, in which the VOC levels are higher than irritation thresholds, therefore, a deeper study of the non-reactive VOCs – sensory irritation relationship has to be carried out. In addition, reactions between certain unsaturated VOCs (e.g. terpenes) and oxidants like environmental ozone produce airway irritants, this phenomenon could possibly explain eye complaints of humans [18]. Reactive oxygenated VOCs such as formaldehyde, acrolein, among other species are formed from these atmospheric reactions, which can occur within the time-frame of low air exchange rates [19]. Since nasal and eye irritation thresholds are comparable [20], it is believed that these observations are also valid for eye irritation, i.e. if chemically reactive VOCs react with ozone they may cause eye irritation as the eye irritation threshold values of the reaction products are much lower.

On the other hand, eye irritation can also be produced through the large number of irritant substances present in the particles from a puff of smoke; these irritants include heavy metals and toxic mineral elements (mercury, lead, cadmium, etc), which are poisonous in high concentrations [21]. Moreover, there are other types of substance present in tobacco smoke such as fume particles and noxious gases with well recognized irritant properties on conjunctival membranes. Another group of components present in tobacco smoke are those with oxidative properties; they are related to some of the most common eye affections. Furthermore, there is another group of components,

photochemical oxidants, which are directly associated with eye disorder and eye irritation [22].

As mentioned in Chapter 2, chemesthesis in the face mucosae (ocular, nasal, and oral) is principally mediated by the trigeminal nerve (cranial nerve V). In the ocular mucosa, chemesthetic stimulation with vapours typically results in eye irritation. Studies at the molecular level have implicated a number of receptor candidates for chemesthesis, in particular, transient receptor potential (TRP) channels but also G protein-coupled receptors (GPCRs) and receptors for chemicals released due to epithelial cell damage such as K^+, H^+, ATP, and glutamate [23,24,25,26]. This broad variety of reception processes agrees with the observation that vapours of volatile organic compounds (VOCs) from very diverse chemical functionality and structure can evoke eye irritation and nasal pungency [9]. Many of these VOCs are relatively non-reactive [27] and, thus, unlikely to damage mucosal tissue, producing the release of the above mentioned mediators, simply upon a brief vapour exposure.

There is evidence that typical pungent compounds other than capsaicin, and structurally very different from it, also activate ion channels directly [28,29,30]. Electrophysiological recordings revealed that multiple types of trigeminal sensory neurons respond differentially to irritating VOCs, suggesting the involvement of multiple mechanisms, some known others yet to be identified, distributed across different modalities of neurons [31]. Within this context, a likely possibility is that the integrated process [32] that governs the chemesthetic potency of most VOC vapours rests heavily on "selective" (or "transfer") effects rather than on "specific" effects.

Support to the above hypothesis came from the successful application of a QSAR that models selective transfer processes across biological matrices [33,34] to human psychophysical thresholds for nasal pungency [35] and eye irritation [36], two typical chemesthetic sensations evoked by VOCs [37]. In turn, the specific equation derived from this model to describe and predict eye irritation thresholds (EIT) is

$$\text{Log}\,(1/\text{EIT}) = -7.918 - 0.482\,\mathbf{E} + 1.420\,\mathbf{S} + 4.025\,\mathbf{A} + 1.219\,\mathbf{B} + 0.853\,\mathbf{L} \qquad (3.1)$$

with N = 54, $R^2 = 0.928$, SD = 0.36, and F = 124, where N is the number of compounds, R is the correlation coefficient, SD is the standard deviation, and F is the F-statistic. Briefly, SP is a property of a series of solutes in a given system, **E** is an excess molar refraction in units of $(cm^3 \cdot mol^{-1}/10)$, **S** is the solute polarisability/dipolarity, **A** and **B** are

the solute hydrogen bond acidity and basicity, respectively. **V** is the McGowan characteristic volume in units of $(cm^3 \cdot mol^{-1})/100$, and **L** is a descriptor that is defined as the logarithm of the equilibrium constant in the solute gas-hexadecane partition coefficient at 25 °C, see Chapter 4 for more details.

Recent studies have provided a detailed description of this QSAR [38], including its specific application to ocular irritation where the model has successfully managed to combine in one equation eye irritation thresholds measured in humans with modified Draize test scores measured in rabbits [39].

$$SP = -7.892 - 0.379\,\mathbf{E} + 1.872\,\mathbf{S} + 3.776\,\mathbf{A} + 1.169\,\mathbf{B} + 0.785\,\mathbf{L} + 0.568\,\mathbf{I} \qquad (3.2)$$

with N = 91, $R^2 = 0.936$, SD = 0.433 and F = 204.5, where again all letters and symbols are as already defined above, except **I** which is an indicator variable that takes the value 1 for the series compounds for which the Draize eye score (MMAS) is obtained, and the value 0 for the series of compounds for which EIT is obtained. Therefore, SP could be the log $(MMAS/P^o)$, where P^o is the saturated vapour pressure, or the log (1/EIT) [39].

3.2. ANATOMY AND PHYSIOLOGY OF THE EYE

To avoid eye irritation and other kind of eye disorders, the human eye has developed some mechanisms of defense. These mechanisms avoid the presence of damaging substances into the eye structure, these substances can be microorganisms which rarely becomes to colonize or infect the ocular surface. One of these mechanisms is the lacrimal secretion. The tears secreted by lacrimal glands have an aqueous nature. The phenomenon of blinking lets the tears distribute over the ocular surface, drain into the lacrimal punctum, and evaporate into the air [40]. Aqueous tears are covered by tear lipid secreted by the meibomian glands that spreads to form an oily layer of pre-ocular tear film.

The tear film is, thus, a pre-ocular film which is not part of the eye but it is intimately associated with it. The thickness of the pre-ocular tear film is approximately 7μm, see Figure 3.1, and it is composed of three main layers: (1) a thin (0.1 μm) anterior lipid layer derived from meibomian gland secretions, (2) a thick (6 μm or more) aqueous layer, and (3) a very thin (0.04 μm) mucinous layer.

The above account is of the tear film structure proposed by Korb and co-workers [41]. Some groups [42,43] have suggested that there is not an accurate knowledge of the thickness of the tear film. As determined by these studies the thickness varies from approximately 3μm to 45 μm, depending on the technique used, that includes invasive and noninvasive techniques.

a) Lipid layer

The *Lipid layer* is composed of a polar and a nonpolar phase. The nonpolar phase is related to an appropriated structure of the polar phase. Whereas the nonpolar lipid phase includes wax, hydrocarbons, cholesterol esters, and triglycerides, the composition of the polar lipid phase includes phospholipids and cerebrosides [44]. The main functions of the lipid layer include: (a) retard evaporation from pre-ocular tear film, (b) it has to be capable of compression and expansion as the eye blinks, (c) the polar phase has to anchor the aqueous layer to the nonpolar phase of the lipid layer.

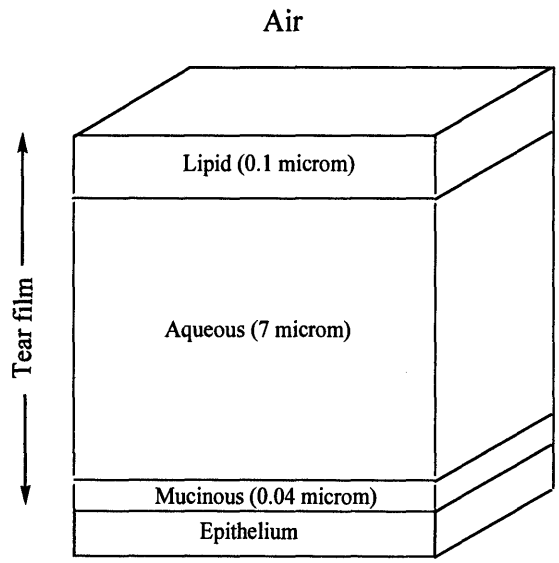


Figure 3.1. Structure of the pre-ocular tear film of the human eye. Source: Ref. 41

b) Aqueous layer

The *Aqueous layer* is a complex mixture of different chemicals, such as lysozyme, albumin, glycoproteins, different types of immunoglobulins, lactoferrin, transferring, ceruloplasmin, etc [45,46]. Furthermore, mucins can be found in this phase. Approximately 98% of the aqueous layer is water whereas 2% is solids, most of them are proteins [45].

c) Mucus layer

The *Mucus layer* is the innermost layer of the tear film and it is composed of mucins, which are secreted mostly by conjunctival goblet cells; the corneal and conjunctival epithelium also contribute to form the mucins of the mucus layer. Mucus is a substance 80-90% hydrated, which is composed of immunoglobulins, urea, salts, glucose, leukocytes, cellular debris and enzymes [47]. Mucins are the most abundant macromolecules in mucus, they are glycoproteins which consist of a protein core, oligosaccharide side chains, and nonglycosylated peptides linked to a glycosylated portion by disulfide bonds (data reported for animals) [48,49]. Dartt *et al.* [50] established that sensory nerves innervate the conjuctiva surrounding goblet cells in rats. So, the deliver of sensory stimuli to the mucus layer of the tear film will lead to a sensory response by the nerves located in this area.

The anchor between the tear film and the cornea is the glycocalyx, the mucus layer is attached to the corneal epithelium by the glycocalyx, this attachment is not tight, and thus, the mucus layer undergoes free movement across the cornea [51]. The supporting glycocalyx of the mucus layer is composed of glycoproteins and glycolipids.

The main functions of the tear film include [52]: (a) maintaining a smooth surface for light refraction, (b) lubricating the eyelids, (c) lubricating the conjunctiva and the cornea, (d) supplying the cornea with nutrients and transporting metabolic by-products from the corneal surface, (e) providing white blood cells with access to the cornea and conjunctiva, (f) removing foreign materials from the cornea and conjunctiva, (g) defending the ocular surface from pathogens via specific and nonspecific antibacterial substances.

The epithelium of the eye is a major protective barrier that regulates transport across the tear film and the inner eye, by making a "tight" sheet with so-called tight junctions.

The nervous structure of the human body is a very complex system. Human body responds to stimuli and the effect is produced because the stimuli are transmitted by the nerves which are composed of fibers. The trigeminal nerve supplies the most important sensory organ (eye, nose and skin of the face) with two types of fibers, myelinated *A-delta fibers* and unmyelinated *C-fibers*. This nerve is sensitive to different types of stimuli, such as pain, temperature, pressure changes, etc. Depending on the type of stimuli, either *A-fibers* or *C-fibers* respond to the stimuli. *A-fibers* can be activated by chemical and thermal stimuli, whilst *C-fibers* respond to intense thermal and mechanical

stimuli. *C-fibers* are the type of fibers that are excited when the sensory organ detects the presence of an irritant component in the environment and both *A-fibers* and *C-fibers* are present in the cornea [53]. The irritation may follow from stimulation of chemical receptors both directly by absorbed pollutants and indirectly, by compounds released in the tissue as a consequence of pollutant exposure [54]. The nerve endings in the eye are located just below the surface of the epithelium, which is protected by the mucus layer. Therefore, any rupture of the tear film will facilitate exposure of the nerve endings to pollutants.

The nature of the processes of irritation by chemicals is barely known; however, some studies have proposed hypotheses including the transduction at the surface of cell membranes by irritative chemicals [55]. Compounds that are chemically reactive can produce irritation directly by reacting with a receptor or indirectly by mucosal tissue damage via chemical reaction without the need to interact with any particular receptor [55].

Eye affections have a logical dose-dependence, so that the risk of developing an affection is higher when the concentration of irritant components in the environment is higher. The severity of the affection has the same dose-dependence. However, for many compounds, the detection limits of existing measuring methods are higher than the exposure levels known or expected to cause effects. The parameter mostly used in the laboratory works to evaluate eye irritation is eye blink response. It is assumed that a higher rate of this parameter is because eye irritation is higher as well.

Several factors may influence the irritation perceived by a subject, whereas inter-individual variation in sensitivity to an irritant can be due to physiological factors, age or exposure history. Extreme variations in subjective evaluations of irritancy can result from cognitive and emotional factors that can influence perception in all sensory modalities [56].

3.3. EYE IRRITATION MEASUREMENTS

To obtain stimulus eye irritation from the single chemicals, once again, as for odour and nasal pungency a two- or three-alternative, forced-choice procedure has been employed with presentation of ascending concentrations [57]. Stimuli are presented either via 250 ml polypropylene squeeze bottles, containing 30 ml of solution, or a 1900 ml glass vessel, containing 200 ml of the solution. These bottles are adapted with a 25 ml,

roughly conical measuring chamber that allowed exposure of one eye at a time to an expelled aliquot from the headspace within the bottle [57], see Figure 3.2.

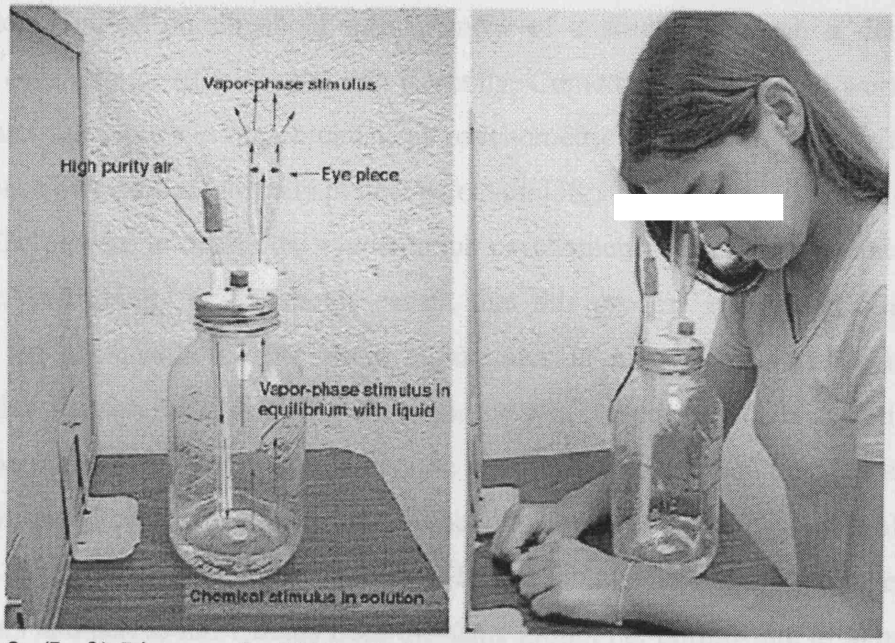


Figure 3.2. (Left) Picture of a 1900 ml glass vessel adapted with an eyepiece for ocular stimulation. (Right) picture of a subject being tested for eye irritation. Source: Ref. 58.

3.3.1 Eye Irritation Detection Threshold

Briefly, the method requires the subject to choose, on each trial, the stronger sensation from the stimuli presented. Unknown to the participant, at least one stimulus is always a blank (mineral oil) and the other a dilution step of the chemical, starting with a step clearly below detection. Over the course of a session, and in ascending order of concentration, each step is presented paired with a blank. If the participant is correct, the same concentration is presented next, paired with blanks. If the participant is incorrect, the next higher concentration is presented next, also paired with blank. The first concentration chosen correctly five times in a row is taken as the threshold. Each eye is tested separately. Each type of threshold is measured eight times per subject–stimulus combination [57]. The ascending concentration order and an interstimulus interval of at least 45 s–60 s helped to minimize any potential changes in trigeminal sensitivity.

3.3.2 Eye Irritation Psychometric Functions

Studies of the functional properties of the trigeminal chemosensory system in human have often focused on threshold measurement of a single point on a detectability function over a range of concentration. Recently, Cometto-Muñiz and co-workers have approached the topic via measurement of psychometric functions spanning the range from chance detection to virtually perfect detection [38,59,60,61].

The process to obtain the eye irritation psychometric functions is similar to that one described above for thresholds except that this process ends when the subject chooses the chemical over the blank eight times in a row, four for each of two consecutive dilution steps [62]. This performance was considered 100% detection. Each subject participated in four sessions like the one described above. In each session, the subject provided complete psychometric function for one chemical. The chemicals and eyes are tested in irregular order. The data from each chemical were averaged within individuals and across individuals from the same group, i.e. normosmic or anosmic. As an example, a psychometric function obtained by Cometto-Muñiz and co-workers for butan-1-ol, using a squeeze bottle to deliver the compound into the eye, is shown in Figure 3.3.

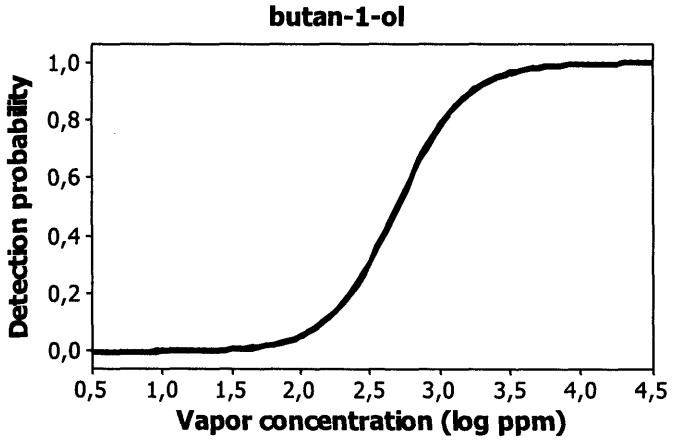


Figure 3.3 Psychometric function for butan-1-ol

3.4. PHYSICOCHEMICAL RESPONSES TO CHEMICAL MIXTURES

As mentioned in section 2.3 of this work, the study of chemical mixtures is attracting considerable attention nowadays, given that many exposures in environmentally

realistic situations involve the simultaneous presence of a number of substances. Previous studies on eye irritation thresholds for mixtures of up to nine VOCs showed agonistic sensory effects among the mixed chemicals [63]. This investigation only measured "thresholds" and did not include complete detectability functions, however it recognized the importance of measuring detectability or concentration-detection (i.e., psychometric) functions for a better characterization of trigeminal detection and a better knowledge of the physicochemical basis that underlies chemical irritation potency.

Later studies with binary mixtures included such functions and found support, as a first approximation, to the notion of chemosensory agonism, in the sense of additivity between the components of the mixtures presented at perithreshold levels [58]. There are indications that the degree of sensory agonism decreases as the detectability of the mixtures approaches high values [59,60,64]. It would be a breakthrough to be able to predict the degree of sensory irritation agonism in mixtures based on the physicochemical and structural properties of the components via a model such as the Abraham Solvation Equation described in Section 3.1 (Eq. 3.1) [36,39]. Nevertheless, such possibility awaits the availability of results from additional and more complex mixtures where the components cover a wide range of properties and structures.

3.5. REFERENCES

1. J. C. Walker, P. R. Nelson, W. S. Cain, M. J. Utell, M. B. Joyce, W. T. Morgan, T. J. Steichen, W. S. Pritchard, M. W. Stancill, Indoor Air 7 (1997) 173
2. A. Bassoli, M. G. Drew, L. Merlini, G. J. Morini, Med. Chem. 45 (2002) 4402
3. B. Bufe, P. A. Breslin, C. Kuhn, D. R. Reed, C. D. Tharp, J. P. Slack, U. K. Kim, D. Drayna, W. Meyerhof, Curr. Biol. 15 (2005) 322
4. M. G. Grigorov, H. Schlichtherle-Cerny, M. Affolter, S. J. Kochhar, Chem. Inf. Comput. Sci. 43 (2003) 1248
5. K. H. Norwich, Chem. Senses 26 (2001) 1015
6. R. A. Sowalsky, A. C. Noble, Chem. Senses 23 (1998) 343
7. P. M. Wise, M. J. Olsson, W. S. Cain, Chem. Senses 25 (2000) 429
8. M. H. Abraham, J. M. R. Gola, J. E. Cometto-Muñiz, W. S. Cain, Chem. Senses 27 (2002) 95

9. J. E. Cometto-Muñiz (2001) Physicochemical basis for odor and irritation potency of VOCs. In J. D. Spengler, J. Samet and J. F. McCarthy (eds.), *Indoor Air Quality Handbook*, McGraw-Hill, New York, p. 20.21
10. B. G. Green, H. T. Lawless (1991) The psychophysics of somatosensory chemoreception in the nose and mouth. In T. V. Getchell, R. L. Doty, L. M. Bartoshuk and J. B. Snow Jr., (eds.) *Smell and Taste in Health and Disease*, Raven Press, New York, p. 235
11. B. G. Green, J. R. Mason, M. R. Kare (1990) Preface. In B. G. Green, J. R. Mason and M. R. Kare, (eds.) *Chemical Senses. Vol. 2: Irritation*, Marcel Dekker, Inc., New York, p. v
12. A. Ferrer-Montiel, C. Garcia-Martinez, C. Morenilla-Palao, N. Garcia-Sanz, A. Fernandez-Carvajal, G. Fernandez-Ballester, R. Planells-Cases, Eur. J. Biochem. 271 (2004) 1820
13. T. Hummel, P. Mohammadian, R. Marchl, G. Kobal, J. Lotsch, Int. J. Psychophysiol. 47 (2003) 147
14. R. L. Doty, J. E. Cometto-Muñiz, A. A. Jalowayski, P. Dalton, M. Kendal-Reed, M. Hodgson, Crit. Rev. Toxicol. 34 (2004) 85
15. C. Fenga, P. Aragona, A. Cacciola, F. Ferreri, G. Spatari, A. Stilo, R. Spinella, D. Germanò, Int. Arch. Occup. Health 74 (2001) 123
16. M. S. Nora, Acta Ophthalmol. Scand. 70 (1992) 269
17. P. Wolkoff, G. D. Nielsen, Atmos. Env. 35 (2001) 4407
18. P. Wolkoff, P. A. Clausen, C. K. Wilkins, G. D. Nielsen, Indoor Air 10 (2000) 82
19. C. J. Weschler, H. C. Shields, Indoor Air 10 (2000) 92
20. J. E. Cometto-Muniz, W. S. Cain, Int. Arch. Occup. Health 71 (1998) 105
21. M. Chiba, R. Masironi, Bull. World Health Organ. 70 (1992) 270
22. J. Schwartz, S. Zeger, Am. Rev. Respir. Dis. 141 (1990) 62
23. D. Julius, A. I. Basbaum, Nature 413 (2001) 203
24. M. Numazaki, M. Tominaga, Curr. Drug Targets CNS Neurol. Disord. 3 (2004) 479
25. J. N. Wood, R. Docherty, Annu. Rev. Physiol. 59 (1997) 457
26. L. Zhang, N. Taylor, Y. Xie, R. Ford, J. Johnson, J. E. Paulsen, B. Bates, Brain Res. Mol. Brain Res. 133 (2005) 187
27. Y. Alarie, G. D. Nielsen, M. H. Abraham, Pharmacol. Toxicol. 83 (1998) 270

28. D. M. Bautista, P. Movahed, A. Hinman, H. E. Axelsson, O. Sterner, E. D. Hogestatt, D. Julius, S. E. Jordt, P. M. Zygmunt, Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 12248
29. S. E. Jordt, D. M. Bautista, H. H. Chuang, D. D. McKemy, P. M. Zygmunt, E. D. Hogestatt, I. D. Meng, D. Julius, Nature 427 (2004) 260
30. L. J. Macpherson, B. H. Geierstanger, V. Viswanath, M. Bandell, S. R. Eid, S. Hwang, A. Patapoutian, Curr. Biol. 15 (2005) 929
31. T. Inoue, B. P. Bryant, Pain 117 (2005) 193
32. M. Tominaga, M. J. Caterina, A. B. Malmberg, T. A. Rosen, H. Gilbert, K. Skinner, B. E. Raumann, A. I. Basbaum, D. Julius, Neuron 21 (1998) 531
33. M. H. Abraham, Chem. Soc. Rev. 22 (1993) 73
34. M. H. Abraham, P. K. Weathersby, J. Pharm. Sci. 83 (1994) 1450
35. M. H. Abraham, R. Kumarsingh, J. E. Cometto-Muñiz, W. S. Cain, Arch. Toxicol. 72 (1998) 227
36. M. H. Abraham, R. Kumarsingh, J. E. Cometto-Muñiz, W. S. Cain, Toxicol. in Vitro 12 (1998) 403
37. M. H. Abraham, J. M. R. Gola, J. E. Cometto-Muñiz, W. S. Cain, Indoor Built Environ. 10 (2001) 252
38. J. E. Cometto-Muñiz, W. S. Cain, M. H. Abraham, Toxicol. Appl. Pharmacol. 207 (2005) 232
39. M. H. Abraham, M. Hassanisadi, M. Jalali-Heravi, T. Ghafourian, W. S. Cain, J. E. Cometto-Muñiz, Toxicol. Sci. 76 (2003) 384
40. F.J. Holly, Am. J. Optom. Physiol. Opt. 57 (1980) 252
41. D. R. Korb, J. Craig, M. Doughty, J-P. Guillon, G. Smith, A. Tomlinson (2002) *The Tear Film Structure, Function and Clinical Examination*, Butterworth-Heinemann, British Contact Lens Association
42. P. E. King-Smith, B. A. Fink, N. Fogt, K. K. Nichols, R.M. Hill, G.S. Wilson, Invest. Ophtalm. Vis. Sci. 41 (2000) 3348
43. J. I. Prydal, P. Artal, H. Woon, F.W. Campbell, Invest. Ophtalm. Vis. Sci. 33 (1992) 2006
44. J. P. McCulley, W. E. Shine, Biosci. Reports 21 (2002) 407
45. S. Iwata, Int. Ophthal. Clinics 13 (1983) 29
46. A. J. German, E. J. Hall, M. J. Day, Vet. Immunol. Immunop. 64 (1998) 107

47. B. A. Nichols, M. L. Chiappino, C. R. Dawson, Invest. Ophthal. Vis. Sci. 26 (1985) 464
48. S. J. Hicks, S. D. Carrington, R. L. Kaswan, S. Adam, J. Bara A. P. Corfieldet, Exper. Eye Research 64 (1997) 597
49. A. J. W. Huang, S. C. G. Tseng, Invest. Ophthalm. Vis. Sci. 28 (1987) 1483
50. D. A. Dartt, D. M. McCarthy, H. J. Mercer, T. L. Kessler, E. H. Chung, J. D. Zieske, Curr. Eye Res. 14 (1995) 993
51. S. Ashutosh, Biophys. Chem. 47 (1993) 87
52. D. W. Lambert (1994) Physiology of the tear film. In G. Smolin, R. A. Thoft (eds.) *The Cornea*, Little Brown & Co, New York, p. 439
53. G.D. Nielsen, Crit. Rev. Toxic. 21 (1991) 183
54. B. Lynn (1984) The detection of injury and tissue damage. In: Wall PD, Melzack R., (eds) *Textbook of Pain*, Churchill Livingstone, Edinburgh, p. 19
55. G.D. Nielsen, Crit. Rev. Toxicol. 21 (1991) 183
56. P. Dalton, Int. Arch. Occup. Environ. Health 75 (2002) 283
57. J. E. Cometto-Muñiz, W. S. Cain, Pharmacol. Biochem. Behav. 39 (1991) 983
58. J. E. Cometto-Muñiz, W. S. Cain, M. H. Abraham, Indoor Air 14 (2004) 108
59. J. E. Cometto-Muñiz, W. S. Cain, M. H. Abraham, J. M. R. Gola, Physiol. Behav. 67 (1999) 269
60. J. E. Cometto-Muñiz, W. S. Cain, M. H. Abraham, J. M. R. Gola, Toxicol. Sci. 63 (2001) 233
61. J. E. Cometto-Muñiz, W. S. Cain, M. H. Abraham, J. M. R. Gola, Appl. Toxicol. 22 (2002) 25
62. J. E. Cometto-Muñiz, W. S. Cain, M. H. Abraham, R. Sanchez-Moreno, Toxicol. Sci. 91 (2006) 600
63. J. E. Cometto-Muñiz, W. S. Cain, H. K. Hudnell, Percept. Psychophys. 59 (1997) 665
64. J. E. Cometto-Muñiz, W. S. Cain, M. H. Abraham, Exp. Brain Res. 158 (2004) 196

Chapter 4 An introduction to the Abraham Solvation Equation

4.0. INTRODUCTION

The purpose of structure activity models such as linear free energy relationship, LFERs, and quantitative structure activity relationships, QSARs, is to quantify the effects of changes in chemical structures on physicochemical activity and biological activity, respectively. First, LFERs are mathematical tools for correlating changes in free energy in different reaction series. The Hammett equation [1,2] is the best known example of a LFER and the work of Hammett has formed a sound basis for many other studies carried out involving LFERs. On the other hand, QSARs, have been used to correlate molecular structural features of compounds with their known biological properties. Attempts to establish relationships between chemical structure and biological effects may be traced back [3] as far as the work of Brown and Fraser [4,5], but Meyer and co-workers [6,7] may have been the first to use a quantitative relationship showing that the product between narcotic concentration and oil/water or solvent/gas concentration ratio is remarkably constant.

The solvation model of Abraham, Kamlet and Taft [8] has offered a good basis for the understanding, description and prediction of the ways in which solutes and solvents interact with each other. One important approach to the problem of solute/solvent interactions has been based on solvatochromism. The solvatochromic comparison approach developed by Kamlet and Taft has provided quantitative scales of solvent hydrogen bond abilities and solvent polarity. A number of quantitative relationships between a variety of physicochemical solute properties and certain solvent parameters have been put forward. Particularly successful and very extensive applications of the solvation model have evolved from the Abraham solvation equation. The Abraham model contains five solute descriptors that are linearly correlated with a given solute property. This model is not simply numerical in nature but offers physicochemical interpretation of a large number of processes. This approach is a well-known and well-used equation for the description of relationships between structure and both physicochemical and biochemical properties.

4.1. LINEAR FREE ENERGY RELATIONSHIP

4.1.1 Solvation Parameter Model

Abraham, Kamlet and Taft developed a simple solvation model [8-10] which describes the solution of a gaseous solute into a solvent. The solution process can be broken down into the following stages:

a) A cavity of suitable size to accommodate the solute is created in the solvent, this process is endoergic because work is required to disrupt solvent-solvent interactions.

b) The solvent molecules around the cavity are reorganized from their original (solvent) position to the positions they will adopt when the solute is in equilibrium with the solvent. The Gibbs energy change for such reorganization is assumed to be negligible.

c) The solute is introduced into the reorganized cavity, and various solute-solvent interactions are set up, all of which are exoergic.

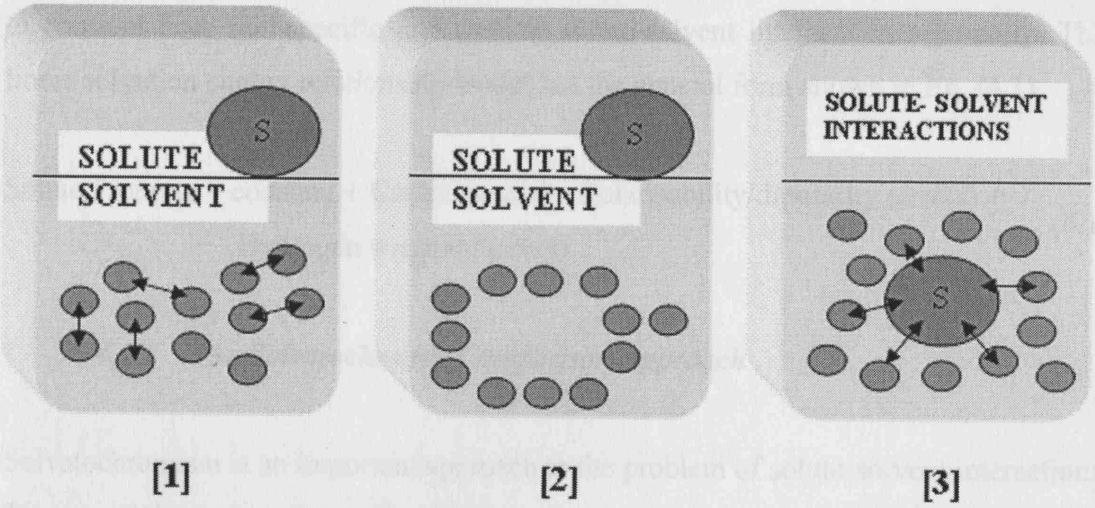


Figure 4.1. The three steps of the cavity model of solvation; ↔ interactions.

The term solvation refers to the surrounding of each dissolved molecule by a shell of more or less tightly bound solvent molecules. This solvent shell is the result of intermolecular forces between solute and solvent [12]. These are called van der Waals forces and are classified in two distinct categories:

- The first category gathers the so-called directional, induction, and dispersion forces, which are non-specific.

- The second category comprises hydrogen bonding forces and the forces of transfer or electron-pair donor/acceptor forces. To this group belong specific, directional forces.

The creation of a cavity is an endoergic effect, as mentioned, the magnitude of which depends on the forces holding the solvent molecules together and the size of the solute. The Gibbs free energy of reorganisation can be taken as zero, and so need to be considered no further. With the introduction of a solute molecule into the solvent cavity, a number of solute-solvent interactions will occur. These interactions are exoergic and aid the processes of solution. Both the cavity term and the solute-solvent interaction term will depend on the properties of the solute and the solvent under consideration. Hence to describe these effects for the general case of a number of solutes in a number of solvents, it is necessary to construct an equation that includes the relevant properties of both the solutes and the solvents.

Abraham, Kamlet and Taft [8,9,12-15] extended the LFERs of earlier workers [16] to involve solute-solvent interactions. Furthermore, they pointed out the necessity to consider both non-specific and specific solute-solvent interactions separately. This linear solvation energy relationship model has the general form shown in Eq. (4.1).

$$\text{Solute Property} = \text{constant} + \text{Cavity term(s)} + \text{Polarisability/dipolarity term(s)} + \text{Hydrogen bonding term(s)} \quad (4.1)$$

4.1.2 The Solvatochromic Comparison Approach

Solvatochromism is an important approach to the problem of solute-solvent interactions. The term solvatochromism is used to describe the pronounced change in position of an UV/Vis absorption band, accompanying a change in the polarity medium. A hypsochromic (or blue) shift, with increasing solvent polarity, is usually called negative solvatochromism. The corresponding bathochromic (or red) shift is termed positive solvatochromism.

Kamlet, Taft *et al.* employed and further developed a solvatochromic comparison method to evaluate a β-scale of solvent hydrogen-bond acceptor (HBA) basicity, an α-scale of solvent hydrogen-bond donor (HBD) acidity, and a π*-scale of

solvent polarisability/dipolarity using UV/Vis spectral data of solvatochromic compounds [17-21]. Magnitudes of enhanced solvatochromic shifts, $\Delta\Delta\tilde{\upsilon}$, in HBA solvents have been determined for 4-nitroaniline relative to homomorphic N,N-diethyl-4-nitroaniline. Both standard compounds are capable of acting as HBA substrates in HBD solvents, but only 4-nitroaniline can act as a HBD substrate in HBA solvents. Taking the $\Delta\Delta\tilde{\upsilon}$-value of 2,800 cm^{-1} for hexamethylphosphoric triamide (a strong HBA) as a single reference point (β_1 = 1.00), a β-scale of solvent HBA basicity for HBA solvents was developed. Using the same solvatochromic comparison method, i.e. the enhanced solvatochromic shift $\Delta\Delta\tilde{\upsilon}$, in HBD solvents for 4-nitroanisole and the pyridinium-N-phenoxide betaine dye, an α-scale of HBD acidity was evaluated. The same authors also introduced a π*-scale of solvent polarisability/dipolarity. The π*-scale is so named because it is derived from solvent effects on the $\pi \rightarrow \pi^*$ electronic transitions of a variety of nitroaromatics. Solvent effects on the $\tilde{\upsilon}_{max}$ values of such solvatochromic indicators have been employed in the initial construction of the π*-scale, which was then expanded and refined by multiple least-squares correlations with additional solvatochromic indicators. In this way an averaged π*-scale of solvent polarisability/dipolarity was established which measures the ability of the solvent to stabilise a charge or a dipole by virtue of its dielectric effect. A normalised range of 0.00 (cyclohexane) to 1.00 (dimethyl sulfoxide) for the π*-values of common solvents was chosen so that, taken with the α-scale of solvent HBD acidity and the β-scale of solvent HBA basicity, these parameters can be used together in a multiparameter equation.

Using the solvatochromic solvent parameters α, β, π*, the multiparameter Eq. (4.2) has been proposed for use in so-called linear solvation energy relationships (LSER) [18-20]. Kamlet, Taft, Abraham *et al.* developed Eq. (4.2) to correlate the solubility and distribution behaviour of non-electrolyte solutes with solvent properties [21,22].

$$SP = SP_0 + A\,\pi_1^*\,\pi_2^* + B\,\alpha_1\beta_2 + C\,\beta_1\alpha_2 + D\,(\delta_H^2)_1\,(V_2/100) \qquad (4.2)$$

SP represents the value of some property in a series of solvents; it can represent, for example, the logarithm of a rate or equilibrium constant, as well as a position of maximal absorption in an UV/Vis, IR spectrum; SP_0 is the regression value of this

solute property in cyclohexane solvent. The subscript 1 refers to the solvent and subscript 2 to the solute. *A, B, C* and *D* are the regression coefficients for the exoergic polarisability/dipolarity term, the exoergic hydrogen bonding terms of adduct formation between HBD solvents and HBA solutes (measured by α_1 and β_2) as well as between HBA solvents and HBD solutes (measure by β_1 and α_2) and the endoergic cavity term, respectively. V_2 is the molar volume of the solute, taken as its molecular mass divided by its liquid density at 293K. The δ_H^2 term represents Hildebrand's solubility parameter squared [23] and corresponds to the cohesive pressure, which characterised the energy associated with the intermolecular solvent-solvent interactions. Thus, δ_H^2 is considered as a measure of the energy required to separate solvent molecules to provide a suitable sized cavity for the solute. Whereas δ^2_H measures the solvent's contribution to the cavity term, $V_2/100$ represents the solute's contribution to the cavity term [24].

When dealing with the effects of different solvents on properties of a single solute, the factors relating to the solute can be subsumed into the regression coefficients of Eq. (4.2), and the following equation results:

$$SP = SP_0 + d\delta + s\pi_1^* + a\alpha_1 + b\beta_1 + h(\delta_H^2)_1 \qquad (4.3)$$

where δ is an empirical polarisability correction term [22] equal to 0.00 for non-chlorine substitued aliphatics, and 0.5 for poly-chlorinesubstituted aliphatics, and 1.0 for aromatic solvents. As mentioned, SP is a property of a given solute in a series of solvents and δ, π_1^*, α_1, β_1, and $(\delta_H)^2$ are the independent variables, that is descriptors of the solvents.

Conversely, Kamlet, Abraham and Taft went back to the full Eq. (4.2), in order to obtain an equation for the case when dealing with solubilities or other properties of a set of different solutes in a single solvent, or with distributions of different solutes between a pair of solvents, the resulting Eq. (4.4) relates the property SP specifically to the solute parameters V_2, π_2^*, α_2, and β_2.

$$SP = SP_0 + s\pi_2^* + d\delta_2 + a\alpha_2 + b\beta_2 + m(V_2/100) \qquad (4.4)$$

The difference between Eq. (4.3) and Eq. (4.4) is very important. Eq. (4.3) refers to the case of a fixed solute, or solutes, in a series of solvents, and is still widely used to

analyse *solvent* effects. Eq. (4.4) refers to the case of a series of solutes in a fixed solvent or solvents, and hence is useful in the analysis of *solute* effects. As will be shown, Eq. (4.4) has now been superseded by the Abraham equation.

One of the earliest tests of Eq. (4.4) was its use in correlating retention in reversed-phase liquid chromatography [26]. Soon thereafter it was used to study octanol-water partition coefficients and solubility in water [27]. Eq. (4.4) works well when applied to processes in condensed phases, but for the processes of the type gas to condensed phase. Abraham *et al.* devised a solute parameter, denoted log L^{16}, where L^{16} is the Ostwald solubility coefficient on n-hexadecane at 298 K. This L^{16} term, as logL^{16}, is related to the endoergic work of creating a cavity in the solvent and the exoergic solute-solvent dispersion interactions [28]. Once added to Eq. (4.4), the term in solute volume was redundant.

$$SP = SP_0 + s\pi_2^* + d\delta_2 + a\alpha_2 + b\beta_2 + l\log L^{16} \quad (4.5)$$

Eqs. (4.4) and (4.5) were revised further by replacing the term δ_2 with R_2, where R_2 is the solute polarisability [29]. In an attempt to overcome some of the pitfalls encountered with the earlier Kamlet and Taft solute parameter [10,11], Abraham *et al.* developed a new polarisability/dipolarity scale, π_2^H and new hydrogen bond acidity and basicity parameters, α_2^H and β_2^H respectively [30-32]. Vx, the McGowan characteristic volume [33,34], was also preferred to V_2. The newly devised descriptors were combined into the two linear equations (4.6) and (4.7) [35].

$$SP = c + rR_2 + s\pi_2^H + a\alpha_2^H + b\beta_2^H + v\mathrm{Vx} \quad (4.6)$$

$$SP = c + rR_2 + s\pi_2^H + a\alpha_2^H + b\beta_2^H + l\log L^{16} \quad (4.7)$$

4.2. THE ABRAHAM SOLVATION EQUATION

4.2.1 The Abraham Solute Parameters

As just mentioned, Abraham *et al.* drew up a list of solute descriptors to characterise solute-solvent interactions occurring during the solvation process. The solute properties or descriptors finally used are given in Table 4.1; both the old complicated notation and

the new notation are shown. Using the simplified notation, the general solvation equations (4.6) and (4.7) become:

$$SP = c + e \cdot \mathbf{E} + s \cdot \mathbf{S} + a \cdot \mathbf{A} + b \cdot \mathbf{B} + v \cdot \mathbf{V} \quad (4.8)$$

$$SP = c + e \cdot \mathbf{E} + s \cdot \mathbf{S} + a \cdot \mathbf{A} + b \cdot \mathbf{B} + l \cdot \mathbf{L} \quad (4.9)$$

where SP is a property of a series of solutes in a given system and where the solute descriptors (**E**, **S**, **A**, **B**, **V** and **L**) are independent variables. **E** is an excess molar refraction in units of ($cm^3 \cdot mol^{-1}/10$) [29], **S** is the solute polarisability/dipolarity [30], **A** and **B** are the solute hydrogen bond acidity [31] and basicity [32], respectively. **V** is the McGowan characteristic volume in units of ($cm^3 \cdot mol^{-1}$)/100 [33,34], and **L** is a descriptor that is defined as the logarithm of the equilibrium constant in the solute gas-hexadecane partition coefficient at 25 °C [28]. The first four descriptors, **E**, **S**, **A**, **B**, can be regarded as measures of the tendency of a solute to undergo various solute-solvent interactions, all of which are energetically favourable, i.e. exoergic. On the other hand, **L** and **V** are both measure of the size of a solute, so will be a measure of the cavity term, see Figure 4.1. Further, since the size of the solute is related to general dispersion interactions, both **L** and **V** also describe solute-solvent dispersion interactions [35].

Table 4.1
Old and new notation of the Abraham solute descriptors

Old Symbol	New Symbol	*Descriptor*
R_2	**E**	Excess molar refraction
π_2^H	**S**	Polarisability/dipolarity
$\Sigma\alpha_2^H$	**A**	Overall solute hydrogen bond acidity
$\Sigma\beta_2^H$	**B**	Overall solute hydrogen bond basicity
$\Sigma\beta_2^0$	$\mathbf{B^0}$	Amended basicity parameter for specific solute/systems
Vx	**V**	McGowan volume
$\log L^{16}$	**L**	Solute gas-hexadecane partition coefficient

The development of the Abraham solvation Eqs. (4.8) and (4.9) to analyse, correlate and predict particular property, SP, requires a knowledge of the relevant Abraham parameters. Thus, the solute parameters need to be identified. They were determined by

Abraham *et al.* by means of a variety of methods mainly based on experimental measurements, and details of how they were initially obtained are covered next. Further in order to overcome the reliance on experimental data, alternative methods to obtain the Abraham solute descriptors have also been proposed and will be discussed later in this chapter. The best advance in this field has been the use of a group contribution approach by Platts *et al.* [36]. This technique allows a rapidly and effective calculation of descriptors.

4.2.1.1. L Parameter

The solute descriptor, **L**, initially formulated by Abraham *et al.* [28], characterises solute size and solute tendency to participate in solute-solvent interactions of the general London dispersion type. **L** is now a well-established descriptor in linear free energy relationship. **L** is defined through L^{16}, Eq. (4.10), the solute Ostwald solubility of a solute in n-hexadecane at 298 K,

$$L^{16} = \frac{\text{concentration of solute in n - hexadecane}}{\text{concentration of solute in the gas phase}} \tag{4.10}$$

Abraham *et al.* chose n-hexadecane as a reference solvent for solute descriptor because n-hexadecane is a readily available non-polar liquid of well-defined structure. **L** values were originally determined by direct GLC measurements on packed columns coated with n-hexadecane at 25 °C [28]. This approach is mainly limited to non-polar and polar volatile solutes at 25 °C, but more importantly interfacial absorption phenomenon contributes non-negligibly to the retention mechanism. So to overcome this drawback, Dallas and Carr [37] have made the use of open tubular fused silica capillary column for which interfacial absorption impact on retention mechanism is small. These direct approaches, however, are restricted to volatile and semi-volatile solutes that have a suitable retention time at 25 °C. The use of predictive models has allowed the determination of **L** values for less volatile solutes. For solutes too involatile at 25 °C, values of **L** can be obtained via GLC measurements on non-polar phases such as squalane or apiezon at elevated temperature [38]. GLC data for a series of solutes can be fitted to an equation of the form,

$$SP = SP_0 + e \cdot \mathbf{E} + l \cdot \mathbf{L} \quad (4.11)$$

Here, SP can be either the logarithm of the retention volume, or relative retention time or can be just the retention index, *I*. Thus, for the large series of solutes studied by Dutoit [32] on a hydrocarbon phase at 110 °C, Abraham *et al.* constructed an equation on the lines of Eq. (4.11). Then further values of **L** can be calculated for any solute for which *I*, or *I*/10, is available. In this way, some 1500 **L** values have been obtained [38].

$$I/10 = 6.669 + 8.918\ \mathbf{S} + 20.002\ \mathbf{L} \quad (4.12)$$
$$n = 138,\ r^2 = 0.9995,\ s = 0.449,\ F = 67950$$

Here and elsewhere, n is the number of data point, r is the correlation coefficient, s is the standard deviation and F is the Fisher statistic.

4.2.1.2. The McGowan Characteristic Volume, V

The cavity term **V** (in $cm^3\ mol^{-1}/100$) [33,34] was chosen by Abraham *et al.* because it can simply be calculated by summation of atomic fragments and number of bonds (B_n) in a molecule (Eq. 2.13), all bonds being counted as one, no matter whether single, double o triple.

$$\mathbf{V} = (\sum \text{atom contributions} - (6.56 \times B_n))/100 \quad (4.13)$$

Some typical values for atom contributions required for the calculation of **V** are given in Table 4.2.

Table 4.2
Atom contributions for calculation of **V** (in $cm^3\ mol^{-1}$)[a]

C = 16.35	P = 24.78	Se = 27.81
N = 14.39	S = 22.91	Br = 26.21
O = 12.43	Cl = 20.95	Sn = 39.35
F = 10.48	B = 18.32	Sb = 37.74
H = 8.71	Ge = 31.02	Te = 36.14
Si = 26.83	As = 29.42	I = 34.35

[a] For each bond between atoms, 6.56 $cm^3\ mol^{-1}$ is to be subtracted.

Actually, it is not necessary to count the number of bonds, B_n, in a complicated molecule because the algorithm of Abraham [35] can be used (Eq. 4.14)

$$B_n = N_a - 1 + R_g \qquad (4.14)$$

where N_a is the total number of atoms and R_g the number of rings.

4.2.1.3. The Solute Hydrogen-Bond Acidity Scale, A

Abraham *et al.* [24] derived a new hydrogen bonding acidity scale, α_2^H, from thermodynamic measurements on 1:1 hydrogen bond complexation, that are related to the Gibbs free energy. The α_2^H parameter is a measure of hydrogen bond acid strength. This scale was drawn from values of equilibrium constants for the 1:1 complexation, log $K_{1:1}$ of acids by reference bases (such as pyridine) in an inert solvent (tetrachloromethane) at 25 °C.

$$A\text{–}H + B \rightleftharpoons A\text{–}H\text{-----}B \qquad (4.15)$$

The hydrogen bond acids and reference bases were present in low concentration so that in solution they were monomeric, nonassociated solutes. Abraham and co-workers studied forty five reference bases which yielded forty five equations in which log$K_{1:1}$ is the dependent variable (Eq. 4.16).

$$\log K_{1:1}\text{ (series of acids A against reference base B)} = L_B \log K^H_A + D_B \qquad (4.16)$$

where L_B and D_B characterize the base, and where the log K^H_A values characterize the series of acids.

The general scale of hydrogen-bond acidity was set up by plotting a series of log K values for acids against a given reference base versus a series of log K values for acids against any other reference base, yielding a series of straight lines. The generated lines intersected at a point where log K^H_A = log $K_{1:1}$ = -1.1, where $K_{1:1}$ is expressed in molar concentration units. This point of intersection correspond to the origin of the scale of hydrogen bond acidity and so solutes with no hydrogen bond acidity have log K^H_A equal to -1.1 units. Further, the various log $K_{1:1}$ plots show family-independent

behaviour, so that it was possible to obtain an 'average' hydrogen-bond acidity for solutes in tetrachloromethane, given as log K^H_A. The origin of the scale was shifted to zero and compressed by transforming log K^H_A into a hydrogen bond acidity scale, α_2^H (Eq. 4.15) [31].

$$\alpha_2^H = (\log K^H_A + 1.1)/4.636 \qquad (4.17)$$

The factor 4.636 is an empirical value, used simply to compress the acidity scale into a convenient working range. In bulk solvent the solute can form multiple hydrogen bonds with the surrounding molecules, therefore the 1:1 complexation no longer applies and the α_2^H values may not be relevant. In the event, Abraham found that the α_2^H values could actually be used as the solute overall or effective hydrogen-bond acidity descriptor, **A**, for most mono-acids. To obtain **A** values a preliminary version of Eq. (4.9) was set up using α_2^H as the hydrogen-bond acid descriptor and was applied to various water-solvent partitions (SP is equal to the logarithmic value of water-solvent partition coefficients) [35,40]. The α_2^H descriptor was then modified where necessary, in order to obtain the effective value, **A**. A new set of equations was then constructed, and the same process repeated until a self-consistent set of equations and **A** values was given. Since the solutes in the water-solvent partitions are surrounded by solvent molecules, the overall hydrogen bond acidity scale is the actual scale required. It was observed that values of **A** were constant along any homologous series, except perhaps for the first one or two members, so once a few values are established, values for the rest of the homologous series can be deduced. Multiple hydrogen bonding of a solute with several solvent molecules gives a higher **A** value than α_2^H, obtained from a simple 1:1 complexation constant, see Table 4.3.

It should be noted that the work carried out to obtain an overall **A** scale from the 1:1 α_2^H scale proceeded side-by-side with the calculation of **A** that took place during the **S** calculation, see section 4.2.1.6. Therefore a constant check had to be made on the self-consistency of the derived **A** values.

Finally, it is important to clarify that hydrogen bond acidity, which indicates the ability of a compound to donate a hydrogen bond, is not related to Brønsted acidity of a compound, which, in turns, refers to loss of a proton. This fact is well illustrated in Table 4.3. It can be seen from their **A** values that acetic acid and phenol do have similar

ability to participate to hydrogen bond interaction. However, acetic acid is a typical Brønsted acid as defined by its pKa value [41] of 4.75 at 25 °C. This is not the case for phenol whose pKa value [41] is given as 9.89 at 25 °C.

Table 4.3
A comparison between α_2^H and **A** values

Solute	α_2^H	**A**
n-Heptane	0.00	0.00
Ethanol	0.33	0.37
Pyrrole	0.41	0.41
Water	0.35	0.82
Acetic acid	0.55	0.61
Phenol	0.60	0.60

4.2.1.4. The Solute Hydrogen-Bond Basicity Scale, **B**

In an exactly similar way, Abraham *et al.* [32,42] established a new hydrogen-bond basicity scale, β_2^H, for solutes using 1:1 hydrogen bond complexation equilibrium constants in tetrachloromethane. log $K_{1:1}$ values for thirty four bases against a given reference acid were used to construct a scale of solute hydrogen bond basicity similar to the solute hydrogen bond acidity scale. The thirty-four equations in terms of log K were of the form,

$$\log K_{1:1} \text{ (series of bases B against reference acid A)} = L_A \log K^H_B + D_A \qquad (4.18)$$

where L_A and D_A are now characteristics of the acid and log K^H_B describes the hydrogen bond basicity of a series of bases. These equations also gave straight lines passing through log K^H_B = log $K_{1:1}$ = -1.1. The hydrogen bond basicity was defined by Eq. (4.19), where the factor 4.636 was chosen so that $\beta_2^H = 1$ for the strong hydrogen bond base hexamethylphosphotriamide.

$$\beta_2^H = (\log K_B^H + 1.1)/4.636 \qquad (4.19)$$

The β_2^H descriptor can be used as the overall or effective hydrogen-bond basicity descriptor, *B*, for mono-bases with a few exceptions. A large number of **B** values were determined in a similar way as **A** [35,40,42,43]. A preliminary version of Eq. (4.9) was

set up using β_2^H as the hydrogen-bond base descriptor and was applied to various water-solvent partitions. The β_2^H descriptor was then modified where necessary, in order to obtain **B** values. A new set of equations was then constructed, and the same process repeated until a self-consistent set of equations and **B** values was given. Table 4.4 gives a comparison between β_2^H and **B** values.

Table 4.4
Comparison between β_2^H and **B** values

Solute	β_2^H	**B**
n-Heptane	0.00	0.00
Diethyl ether	0.45	0.45
Butanone	0.18	0.51
Acetonitrile	0.44	0.32
Ethanol	0.44	0.48
Benzene	0.15	0.14
Pyridine	0.62	0.52
Dimethyl sulfoxide	0.78	0.88
Phenol	0.22	0.30
4-Methoxyaniline	0.45	0.65
1,4-Dioxane	0.47	0.64

Abraham introduced an additional hydrogen bond basicity term, $\mathbf{B^0}$, for solutes such as sulfoxides, anilines, pyridines, and some heterocyclic compounds in water-solvent partitioning systems where the solvent phase is quite aqueous. The latter include n-octanol, ethyl acetate, n-butyl acetate, diethyl ether and di-n-butyl ether. For non-aqueous phases such as chloroform, alkanes, benzene and the gas phase, the original **B** term can be used for all solutes.

4.2.1.5. The Excess Molar Refraction, E

The polarisability-correction descriptor, δ_2, used in Eq. (4.2), is only an empirical factor limited to one of three values, 0.5 for halogenated aliphatics, 1.0 for aromatics or 0.0 for all other compounds. A number of possible alternatives for δ_2 were considered by

Abraham *et al.* [29], molar refraction (MR_X) being one of them. Molar refraction is often used as a measure of polarisability and can be defined as:

$$MR_X = 10[(n^2-1)/(n^2 + 2)]\mathbf{V} \quad (4.20)$$

Here n is the refractive index of a solute that is liquid at 20°C (for solids, the refractive index of the hypothetical liquid at 20°C can be calculated) and **V** is the McGowan characteristic volume in $(cm^3\ mol^{-1})/100$ [33,34]. Because of the volume term in molar refraction, the latter always increases with increasing size. The refractive index function itself is rather better indication of the presence of polarisable electrons in a molecule; thus values of the refractive index are always larger for aromatic or halogenated aliphatic compounds than for other aliphatics.

The molar refraction of a solute in 'excess' of the molar refraction of an alkane of the same characteristic volume can be defined as **E** ($10^{-1}cm^3\ mol^{-1}$), where:

$$\mathbf{E} = MR_X(\text{observed}) - MR_X(\text{for alkane of the same } \mathbf{V}) \quad (4.21)$$

By substracting the molar refraction for an alkane of the same characteristic volume, the dispersive component of molar refraction (already accounted in **V** and **L** in LFER correlations) is removed. **E** provides a quantitative indication of polarisable n and π electrons. **E** is an almost additive quantity that can easily be estimated for solids, and for structures in general, from fragment or substructure values [33,35].

4.2.1.6. The Solute Polarisability/Dipolarity Scale, S

Originally π_2^* was taken as the solvent parameter π_1^* for non-associated liquids set out by Kamlet and Taft [8,11,20,44,45]. As π_1^* is experimentally accessible only for compounds that are liquid at 25 °C, values of π_2^* had to be estimated for associated compounds such as acids, phenols, alcohols and amides as well as gaseous and solid solutes. In addition, there is present the inherent assumption that π_1^* is identical to π_2^* for non-associated liquids, but this may not always be the case. Furthermore, because of its spectroscopic origin, this parameter fails to be Gibbs energy related. It therefore seemed necessary to develop a method that would allow the determination of a

polarisability/dipolarity scale, **S**, that would be free energy related and include all types of solute molecule.

Abraham *et al.* constructed the new polarisability/dipolarity parameter, **S**, from the extensive sets of retention gas liquid chromatographic (GLC) data. The use of the McReynolds [46] and Patte *et al.* [47] chromatographic data provided **S** values for hundreds of solutes. **S** values for substituted aromatics, polyhalogenated aliphatics, nitroalkanes and nitriles were obtained using the general solvation equation on Fellous *et al.* retention data for 17 stationary phases [48]. The **S** values for halogenated or polyhalogenated solutes were again obtained by the same method using retention data for various other stationary phases [49,50].

4.2.1.7. Advances in the Abraham Solute Descriptor Determination

a) Estimation of Solute Descriptors from Experimental Data

The **V** descriptor can always be determined from the solute structure. Most of the time, **E** is easily calculated. In such a case, **S**, **A**, **B** and **L** remain to be determined. Abraham *et al.* developed a general procedure to simultaneously determine descriptor values. The method is based on the use of Eqs. (4.8) and (4.9) applied to as many physicochemical properties as possible. In Table 4.5, are listed some typical physicochemical properties in use.

Table 4.5
Water-solvent and gas-solvent processes used in the determination of solute descriptors.

Systems	Dependent Variable	SP
Water-solvent partition	P: water/solvent partition coefficient	log P
HPLC	k': capacity factor	log k'
Gas-solvent partition	L: gas/solvent partition coefficient	log L
GLC	I: retention index	I
	t_{rel}: relative retention time	log t_{rel}
	Vg: specific retention volume	log Vg

The descriptor values for **S**, **A**, **B** and **L** are taken as the most statistically sound descriptor values that satisfy the various physicochemical properties, already calibrated through known solvation equations. In this way, a database of descriptors for some 3500

compounds has been established, see Table 4.6. The general method has been recently detailed for the determination of descriptors for terpenes [52] and buckminsterfullerene [53].

Table 4.6
Table of the available solute descriptors

Descriptor	Maximum value	Minimum value	Total
E	4.62	-1.21	5957
S	5.70	-1.80	5501
A	4.33	0.00	6137
B	5.68	0.00	5233
V	8.56	0.07	6194
L	34.48	-1.74	4349

b) Estimation of Solute Descriptors from Empirical Methods

Whilst methods based on experimental data deliver descriptors for most molecules, a number of drawbacks exist. First, one must physically obtain a sample of the solute of interest. Second, certain measurements may not be suited to certain types of solute. Third, the process of measurement is time-consuming, and laborious. Finally, this approach, based on experimental measurements, limits the possibility of using it in so-called high-throughput screening, the rapid evaluation of molecular properties for large libraries of compounds [36]. Consequently, a number of methods for the estimation of solute descriptors that do not require experimental data but that are often based on computed quantities, have been proposed. The various approaches are now presented, attention is principally drawn to the group contribution approach used by Platts *et al.* [36].

- Estimation from Structure

In any homologous series, the value of **S**, **A** and **B** remain almost constant, apart form the first two members. Thus, homologous are dealt with ease. For branched compounds, the values of **S** often decreases by 0.03 units for each branch, compared to the unbranched compounds, this is particularly applicable to aliphatic compounds. **L** values along homologous series are well correlated with carbon number, N_C, and for any such

series a plot of **L** versus N_C will yield to a good regression equation, from which further **L** values can be extrapolated [38].

- Estimation from Solute Physicochemical Properties

Abraham *et al.* [31] put forward a reasonable correlation (Eq. 4.22) between the solute hydrogen bond acidity, α_2^H, and the Hammett inductive parameter, σ_I, for a few halogenated compounds, such as chloroform, dichloroethane, 1-chloro-1,1,2-trifluoro-2-iodoethane, α,α-dibromotoluene, methyl dichloroacetate and 1,1,1-trifluoro-2-bromo-2-chloroethane.

$$\alpha_2^H = -0.114 + 0.992\,\sigma_I \qquad (4.22)$$

$n = 18,\ r = 0.91,\ s = 0.02$

This equation is good enough to calculate additional α_2^H values for similar halogenated solutes.

Abraham [50] proposed the estimation of **S** values for chlorinated benzenes (Eq. 4.23) through the use of solute dipole moment, μ, and the number of chlorine atoms, N_{Cl}. The choice of these two solute properties was driven by the interest in dissecting **S** values into contributions from dipolarity (μ) and polarisability (N_{Cl})

$$\mathbf{S} = 0.538 + 0.0743\,N_{Cl} + 0.0353\mu \qquad (4.23)$$

$n = 13,\ r = 0.98,\ s = 0.03$

- Estimation from Quantum Properties

Murray and Politzer [54] have developed a general approach that permits the analysis, correlation and prediction of Abraham hydrogen bonding parameters from a series of computed quantities evaluated on solute molecular surface. The authors have shown that there is a reasonable relationship between the calculated surface maxima, $V_{S,max}$, that describes the electrostatic potential associated with hydrogen atoms in the solute, and the descriptors α_2^H and **A**, see Eqs. (4.24) and (4.25).

$$\alpha_2^H = -0.371 + 0.0257\ V_{S,max} \quad (4.24)$$

$n = 15, r = 0.9685, s = 0.05, F = 199.6$

$$\mathbf{A} = -0.316 + 0.0246\ V_{S,max} \quad (4.25)$$

$n = 15, r = 0.9731, s = 0.04, F = 222.4$

Since the usual error in hydrogen bond parameters is around 0.03 units, it is possible to estimate further **A** values, at least for monofunctional oxygen acids. Similarly, Murray and Politzer [54] have shown a sufficient correlation between the electrostatic potential minimum, V_{min}, and β_2^H for a series of oxygen bases and a series of nitrogen bases. The oxygen and nitrogen bases have to be taken separately. Here is an example for the oxygen bases

$$\beta_2^H = -0.228 - 0.0134\ V_{min} \quad (4.26)$$

$n = 16, r = 0.9554, s = 0.065, F = 146.4$

The above equation can be used to calculate β_2^H values for oxygen bases. The authors have also tried to establish a similar relationship using the overall basicity descriptors, **A** and $\mathbf{A^0}$, but the results showed that these descriptors do not correlate well to V_{min} [54].

Sevcik and co-workers [55] have made use of a neural network approach to estimate **S** values. The authors took a number of structural and quantum mechanical properties as input, combining them either linearly via multivariate linear regression analysis (MLRA) or nonlinearly via a feed-forward neural network.

- Estimation from Group Contribution Methods

In group contribution approach, molecules are broken down to predefined fragments and their corresponding contributions are summed up to obtain the final descriptor values. Molecules however are never mere collection of fragments. Group contribution methods attempt to account for this by introducing different correction factors that are also considered additive. Consequently, all such methods rely on a basic equation of the type

$$\text{descriptor value} = \sum_{i=1}^{n} a_i f_i + \sum_{j=1}^{m} b_j F_j \qquad (4.27)$$

where f_i represents the fragmental contribution and a_i represents the number of occurrences of fragment type i, F_j represents the contribution and b_j represents the number of occurrences of correction factor j [56].

The group contribution method was first applied to estimation of the Abraham descriptors by Sevcik *et al.*[55]. Their approach simply consists on adding contributions to L from a given set of fragments, the contribution being derived by MLRA. Recently, Platts *et al.* [36] have developed additive models for the estimation of Abraham's molecular descriptors **E**, **S**, **A**, **B**, $\mathbf{B^0}$, and **L**. From a database of between 2500 and 3500 values for each descriptor, Platts *et al.* were able to identify common substructures and, through a process of MLRA they evaluated contributions of each substructures to each descriptors. Their final model used 81 atoms and functional group fragments for **E**, **S**, **A**, **B**, $\mathbf{B^0}$, and **L** and was able to reproduce experimentally derived results with correlation ranging from 0.95 to 0.99. However, **A** required an entirely separate set of 51 fragments to be developed, resulting in a correlation coefficient of 0.97. Typically, errors of around 0.05-0.15 log unit (for values covering a range of 2-6 log units) were found [36].

Of particular importance is the speed of calculation allowing rapid evaluation of molecular properties for large libraries of compounds. Once a list of SMILES strings is entered in the program, the descriptors can be calculated for up to 50 molecules per minute in a PC and for up to 700 molecules per minute using a UNIX version. When the descriptors are known, computer calculation of properties from the regression equations of type of Eqs. (4.8) and (4.9), is trivial, and so a number of properties of molecules can be estimated very rapidly from structure. This model has been trained and applied to several systems [36,57,58].

A software package, ABSOLV, based on a group contribution approach similar to that developed by Platts *et al.*, is now commercially available. ABSOLV uses the simplified Abraham solute descriptor notation, see Table 4.1 [59]. Another software package, recently developed by Pharma Algorithms, is ADME Boxes, of which the latest version v3.2 has been released in 2006. This package, which supersedes the ABSOLV calculation, also uses a group contribution approach to calculate the Abraham descriptors, and can be accessed just by introducing the SMILES notation for

compounds.

4.2.2 Applications of the Abraham Solvation Equation

The Abraham general solvation equations are examples of Linear Free Energy Relationships (LFERs) that correlate a physical or chemical property for a set of solutes with a corresponding set of solute physicochemical property descriptors (**E**, **S**, **A**, **B** and **V/L**). When applied to biological properties, the general equations will then refered to Quantitative Structure Activity Relationships (QSARs). The equation coefficients (*c*, *e*, *s*, *a*, *b* and *v/l*) obtained are dependent on the system under investigation and can be used to predict or estimate further values of the independent variable for a completely new solute providing the descriptors are known. In addition, the equation coefficients provide information on the phase system. For a partition between two condensed phases, Eq. (4.8) is used, equation coefficients will then refer to differences in physicochemical properties of the two phases. The *e* coefficient is a measure of difference in phase polarisability and the *s* coefficient is a measure of phase polarisability/dipolarity difference. The *a* coefficient measures the difference in the two phases hydrogen-bond basicity (because an acidic solute will interact with a basic phase) and the *b* coefficient is a measure of how the phases differ in hydrogen-bond acidity. The *v* coefficient is a combination of exoergic dispersion forces that make a positive contribution to the *v* coefficient and an endoergic cavity term that makes a negative contribution. The dispersion interaction nearly always dominates so that the *v* coefficient is usually positive except for solution of gases and vapours in water. The *v* coefficient is a useful measure of how the hydrophobicity of the two phases differs. Eq. (4.9) is simpler and is applied to processed involving gas to condensed phase transfer. The *l* coefficient is also resultant of dispersion and cavity effects and is usually positive. Since the *l* coefficient varies between -0.21 for water at 25 °C and +1.00 for n-hexadecane at 25 °C, it seems to be a suitable measure of condensed phase lipophilicity [38].

It is important to note that for gas to condensed phase processes the *s*, *a* and *b* coefficients must always be positive, or zero, because interactions occurring between a condensed phase and a solute must increase the solubility of the solute. The *e* coefficient is an exception because it is tied to hydrocarbons as zero; hence phases containing fluorinated or chlorinated compounds may give rise to a negative *e*

coefficient. Therefore, the coefficients in the solvation parameter equation are not just fitting constants but must obey general chemical principles [38].

The generality of the solvation equations is highlighted by the fact that they have been applied to a hugely diverse range of processes. Eq. (4.8) has been employed for processes that take place in condensed phases, such as water-solvent partitions [59,60], water-micelle partitions [62], high performance liquid chromatography (HPLC) [63], normal phase liquid chromatography [64], microemulsion [65] and micellar [66] electro-kinetic chromatography, thin-layer chromatography [67], solid phase extraction [68], blood-brain distribution [69,70], brain perfusion [71], water-skin permeation [70,64], and tadpole narcosis [73]. Abraham and Le [74] have recently shown that a modified form of the Abraham solvation equation can be used to calculate and predict the solubility of solids and liquids in water. Eq. (4.8) has been applied to the prediction of gastrointestinal absorption values [75] as well. Eq. (4.9) has been applied to a numerous gas-solvent partitions [61], to gas-biological phase partitions [76], and to a very large number of gas chromatographic systems [59]. Similarly, this equation has been used to correlate odor detection threshold [77], nasal pungency threshold [78], eye irritation threshold [78], and to predict respiratory irritation in mice [79]. Both Eqs. (4.8) and (4.9) are now well tried and tested equations.

4.3. REFERENCES

1. N. S. Isaacs, (1995) in *Physical Organic Chemistry*, 2nd edn. Longman Scientific and Technical, Harlow
2. L. P. Hammett, J. Am. Chem. Soc. 59 (1937) 96
3. W. P. Purcell, G. E. Bass, J. M. Clayton, (1973) in *Strategy to Drug Design: a guide to Biological Activity*, Wiley Interscience: New York
4. C. B. Fraser, T. R. Fraser, Proc. Roy. Soc. Edinburgh, 6 (1867-1868) 228
5. C. B. Fraser, T. R. Fraser, Proc. Roy. Soc. Edinburgh, 6 (1867-1868) 461
6. K. H. Meyer, H. Gottlieb-Billroth, Physiol. Chem. 112 (1920) 55
7. K. H. Meyer, H. Hemmi, Biochem. Z. 277 (1935) 39
8. M. J. Kamlet, R. M. Doherty, J.-L.M. Abboud, M. H. Abraham, R. W. Taft, Chem. Tech. (1986) 566
9. M. H. Abraham, R. M. Doherty, M. J. Kamlet, R. W. Taft, Chem. Br. 22 (1986) 551

10. M. J. Kamlet, R. M. Doherty, J.-L. M. Abboud, M. H. Abraham, R. W. Taft, J. Pharm. Sci. 75 (1986) 338
11. M. J. Kamlet, R. M. Doherty, M. H. Abraham, P. W. Carr, R. F. Doherty, R. W. Taft, J. Phys. Chem. 91 (1987) 1996
12. C. Reichardt, (1990) in *Solvents and Solvent effects in Organic Chemistry*, second, revised and enlarged edition, VCH-Weinheim
13. M. J. Kamlet, R.W. Taft, J.-L. M. Abboud, J. Am. Chem. Soc. 99 (1977) 8325
14. M. H. Abraham, R. M. Doherty, M. J. Kamlet, R. W. Taft, Chem. Br. 22 (1986) 551
15. M. J. Kamlet, R. M. Doherty, M. H. Abraham, P. W. Carr, R. F. Doherty, R. W. Taft, J. Phys. Chem. 91 (1987) 1996
16. I. A. Koppel, V. A. Palm, (1972) in *Advances in Linear Free Energy Relationships*, N.B. Chapman and J. Shorter (eds.), Plenum Press, London
17. M. J. Kamlet, R.W. Taft, J. Am. Chem. Soc. 98 (1976) 2886
18. M. J. Kamlet, J.-L. M. Abboud, R. W. Taft, J. Am. Chem. Soc. 99 (1977) 6027
19. J.-L. M. Abboud, M. J. Kamlet, R. W. Taft, Progr. Phys. Org. Chem. 13 (1981) 485
20. R.W. Taft, J.-L. M. Abboud, M. J. Kamlet, M. H. Abraham, J. Solution Chem. 14 (1985) 153
21. M. J. Kamlet, R.W. Taft, Acta Chem. Scand., Part B 40 (1986) 619
22. M. J. Kamlet, M. H. Abraham, R. M. Doherty, R. W. Taft, J. Am. Chem. Soc. 106 (1984) 464
23. M. J. Kamlet, R. M. Doherty, M. H. Abraham, R. W. Taft, Quant. Struct. Act. Relat. (1988) 71
24. M. J. Kamlet, J.-L. M. Abboud, M. H. Abraham, R. W. Taft, J. Org. Chem. 48 (1983) 2877
25. J. H. Hildebrand, R. L. Scott, (1964) in *The Solubility of Non-Electrolytes*, Dover Pub., N.Y.
26. R. W. Taft, J. S. Murray, (1994) in *Quantitative Treatments of Solute/Solvents Interactions, Theoretical and Computational Chemistry*, Vol.1, P. Politzer, J. S. Murray (eds.), Elsevier Science B.V., p. 55
27. M. J. Kamlet, R. M Doherty, M. H. Abraham, Y. Marcus, R. W. Taft, J. Phys. Chem. 92 (1988) 5244
28. M. H. Abraham, P. L. Grellier, R. A. McGill, J. Chem. Soc., Perkin Trans. 2 (1987) 797

29. M. H. Abraham, G. S. Whiting, R. M. Doherty, W. J. Shuely, J. Chem. Soc. Perkin Trans 2 (1990) 1451
30. M.H. Abraham, G.S. Whiting, R.M. Doherty, W.J. Shuely, J. Chromatogr. A 587 (1991) 213
31. M. H. Abraham, P. L. Grellier, D. V. Prior, P. P. Duce , J. J. Morris, P. J. Taylor, J. Chem. Soc. Perkin Trans 2 (1989) 699
32. M. H. Abraham, P.L. Grellier, D.V. Prior, J. J. Morris, P. J. Taylor, C. Laurence, M. Berthelot, Tetrahedron Lett. 30 (1989) 2571
33. J. C. McGowan, J. Appl. Chem. Biotechnol. 34A (1984) 38
34. M. H. Abraham, J. C. McGowan, Chromatographia 23 (1987) 243
35. M. H. Abraham, Chem. Soc. Revs. 22 (1993) 73
36. J. A. Platts, D. Butina, M.H. Abraham, A. Hersey, J. Chem. Inf. Comput. Sci. 39 (1999) 835
37. A. J. Dallas, P.W. Carr, J. Phys. Chem. 98 (1994) 4927
38. M. H. Abraham, (1994) in *Quantitative Treatments of Solute-Solvent Interactions Theoritical and Computational Chemistry* Vol.1, P. Politzer, J.S. Murray (eds.), Elsevier Sci., B.V., p. 83
39. J.-C. Dutoit, J. Chromatogr. A 555 (1991) 191
40. M. H. Abraham, H. S. Chadha, (1996) in *Applications of a Solvation Equation to Drug Transport Properties, in Lipophilicity in Drug Action and Toxicology*, V. Pliska, B. Testa and H. van de Waterbeemd (eds.), VCH, Weinheim, Germany
41. R. C. Weast, (1974) in *Handbook of Chemistry and Physics*, 55th edn, CRC Press Inc., Cleveland
42. M. H. Abraham, P. L. Grellier, D. V. Prior, J. J. Morris and P. J. Taylor, J. Chem. Soc. Perkin Trans. 2 (1990) 521
43. M. H. Abraham, J. Phys. Org. Chem. 6 (1993) 660
44. M. H. Abraham, M. J. Kamlet, R. W. Taft, R. M. Doherty, Nature (London) 106 (1984) 464
45. R. W. Taft, M. H. Abraham, G. R. Famini, R. M. Doherty, J.-L. M. Abboud, M. J. Kamlet, J. Pharm. Sci. 74 (1985) 807
46. W. O. McReynolds, (1966) in *Gas Chromatographic Retention Data*, Preston Technical Abstracts, Evanston, IL
47. F. Patte, M. Etcheto, P. Laffort, Anal. Chem. 54 (1982) 2239
48. R. Fellous, L. Lizzani-Cuvelier, R. Luft, Anal. Chim. Acta 174 (1985) 53

49. M. H. Abraham, G. S. Whiting, J. Chromatogr. A 594 (1992) 229
50. M. H. Abraham, J. Chromatogr. 644 (1993) 95
51. A. Leo, the Medicinal Chemistry Project Pomona College, Claremont, CA, 91711
52. M. H. Abraham, R. Kumarsingh, J. E. Cometto-Muniz, W. S. Cain, M. Roses, E. Bosch, M. L. Diaz, Arch. Toxicol. 27 (1998) 227
53. M. H. Abraham, C. E. Green, W. E. Acree, Jr., J. Chem. Soc., Perkin Trans. 2 (2000) 281
54. J. S. Murray, P. Politzer, J. Chem. Res. (S) (1992) 110
55. P. Havelec, J. G. K. Sevcik, J. Phys. Chem. Ref. Data 25 (1996) 1483
56. P. Buchwald, N. Bodor, Curr. Med. Chem. 5 (1998) 353
57. J. A. Platts, M. H. Abraham, D. Butina, A. Hersey, J. Chem. Inf. Comput. Sci. 40 (2000) 71
58. J. A. Platts, M. H. Abraham, Environ. Sci. Technol. 34 (2000) 318
59. Absolv, syrius
60. M. H. Abraham, H. S, Chadha, G. S. Whiting, R. C. Mitchell, J. Pharm. Sci. 83 (1994) 1085
61. M. H. Abraham, C. E. Green, W. E. Acree, Jr., C. E. Hernandez, L. E. Roy, J.Chem. Soc., Perkin Trans. 2 (1998) 2677
62. M. H. Abraham, H. S. Chadha, J. P. Dixon, C. Rafols, C. Treiner, J. Chem. Soc., Perkin Trans. 2 (1995) 887
63. M.H. Abraham and M. Roses, J. Phys. Org. Chem. 7 (1994) 672
64. F. Z. Oumada, M. Rosés, E. Bosch, M. H. Abraham, Anal. Chim. Acta 382 (1999) 301
65. M. H. Abraham, C. Treiner, M. Roses, C. Rafols, Y. Ishihama, J. Chromatogr. A 752 (1996) 243
66. S. K. Poole, C. F. Poole, Analyst 122 (1997) 267
67. M. H. Abraham, C. F. Poole, S. K. Poole, J. Chromatogr. A 749 (1996) 201
68. D. S. Siebert, C. F. Poole, M. H. Abraham, Analyst 121 (1996) 511
69. M. H. Abraham, A. Ibrahim, Y. Zhao, W. E. Acree Jr, J. Pharm. Sci. 95 (2006) 2091
70. M. H. Abraham, A. Ibrahim, W. E. Acree Jr., Eur. J. Med. Chem. 41 (2006) 494
71. J. A. Gratton, M. H. Abraham, M. W. Bradbury, H. S. Chadha, J. Pharm. Pharmacol. 49 (1997) 1211
72. M. H. Abraham, H. S. Chadha, R. C. Mitchell, J. Pharm. Pharmacol. 47 (1995) 8
73. M. H. Abraham, C. Rafols, J. Chem. Soc., Perkin Trans.2 (1995) 1843
74. M. H. Abraham, J. Le, J. Pharm. Sci. 88 (1999) 868

75. Y. H. Zhao, J. Le, M. H. Abraham, A. Hersey, P. J. Eddershaw, C. N. Luscombe, D. Boutina, G. Beck, B. Sherbourne, I. Cooper, J. A. Platts, J. Pharm. Sci. 90 (2001) 3484
76. M. H. Abraham, P. K. Weathersby, J. Pharm Sci. 83 (1994) 1450
77. M. H. Abraham, J. M. R. Gola, J. E. Cometto-Muniz, W. S. Cain, Chem. Senses 27 (2002) 95
78. M. H. Abraham, J. M. R. Gola, J. E. Cometto-Muniz, W. S. Cain, Indoor Built Environ. 10 (2001) 252
79. Y. Alarie, M. Shaper, G. D. Nielsen, M. H. Abraham, Arch. Toxicol. 72 (1998) 125

Chapter 5 An Introduction to Gas Liquid Chromatography

5.0. INTRODUCTION

Chromatography is essentially a physical method of separation in which the components to be separated are distributed between two phases, one of which is stationary (stationary phase) while the other (the mobile phase) moves in a definite direction [1,2].

Useful chromatographic separations require an adequate difference in the strength of physical interactions for the sample components in the two phases, combined with a favourable contribution from system transport properties that control the movement within and between phases [3]. Several key factors are responsible, therefore, or act together, to produce an acceptable separation. Individual compounds are distinguished by their ability to participate in common intermolecular interactions in the two phases, which can generally be characterized by an equilibrium constant, and is thus a property predicted from chemical thermodynamics. During transport through or over the stationary phase differential transport resulting from diffusion, convection, turbulence, etc., results in dispersion of solute zones around an average value, such that they occupy a finite distance along the stationary phase in the direction of migration. The extent of dispersion restricts the capacity of the chromatographic system to separate, and independently of favourable thermodynamic contributions to the separation, there are a finite number of dispersed zones that can be accommodated in the separation. Consequently, chromatographic separations depend on a favourable contribution from thermodynamic and kinetic properties of the compounds to be separated [4].

A convenient classification of chromatographic techniques can be made in terms of the physical state of the phases employed for the separation. In this case, when the mobile phase is a gas and the stationary phase is a liquid the separation technique is known as gas-liquid chromatography (GLC). Gas-liquid chromatography is the more popular separation mode and is often simply referred to as gas chromatography (GC).

The principal function of the gas chromatograph is to provide those conditions required by the column for achieving a separation without adversely affecting its performance in any way. Operation of the column requires a regulated flow of carrier

gas, an inlet system to vaporise and mix the sample with the carrier gas, a thermostatted oven to optimise the temperature for the separation, an on-line detector to monitor the separation; and associated electronic components to control and monitor instrument conditions, and to record, manipulate and format the chromatographic data [5]. Figure 5.1 shows a schematic diagram of the principal components of a gas chromatograph.

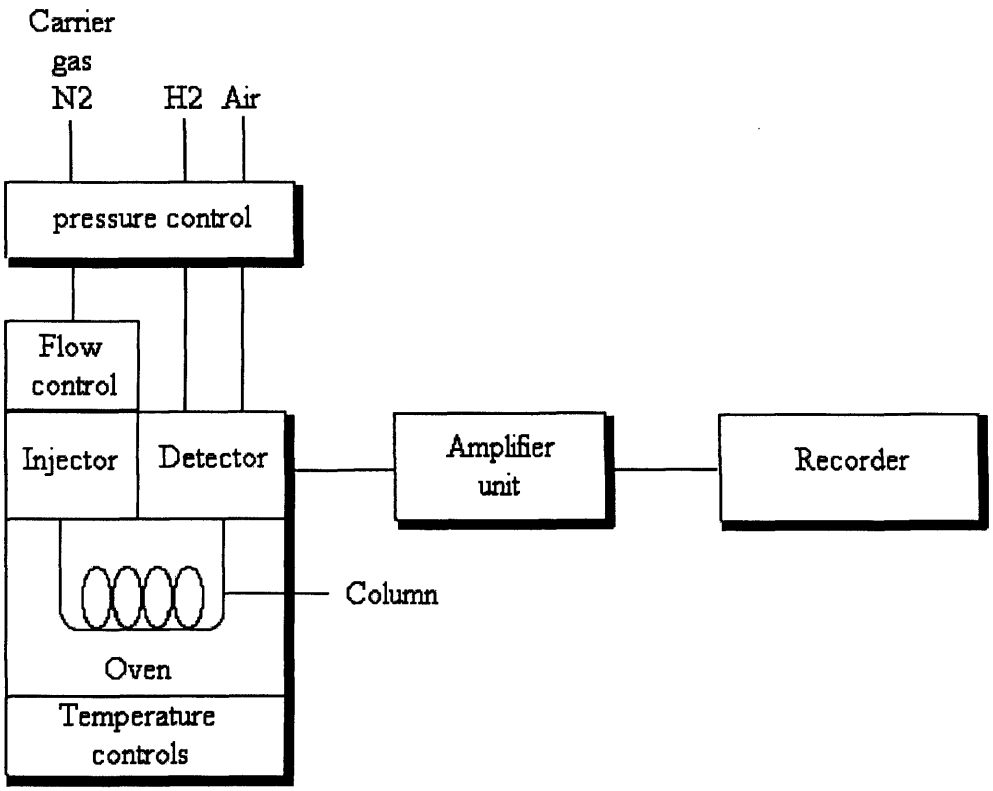


Figure 5.1. Schematic diagram of the principal components of a gas chromatograph.

In elution chromatography, which is the mode of zone displacement used in this work, the mobile and the stationary phase are normally at equilibrium. The sample is applied to the column as a discrete band and sample components are successively eluted from the column diluted by the mobile phase. The mobile phase must compete with the stationary phase for the sample components and for a separation to occur the distribution constants for the sample components resulting from the competition must be different.

The information obtained from a chromatographic experiment is contained in the chromatogram. When the elution mode is used this consists of a plot of (usually)

detector response (y-axis) as a continuous function of time or volume of mobile phase passed through the column (x-axis). The chromatogram contains a number of peaks of various sizes rising from a baseline, Figure 5.2. It is useful for both qualitative and quantitative analysis.

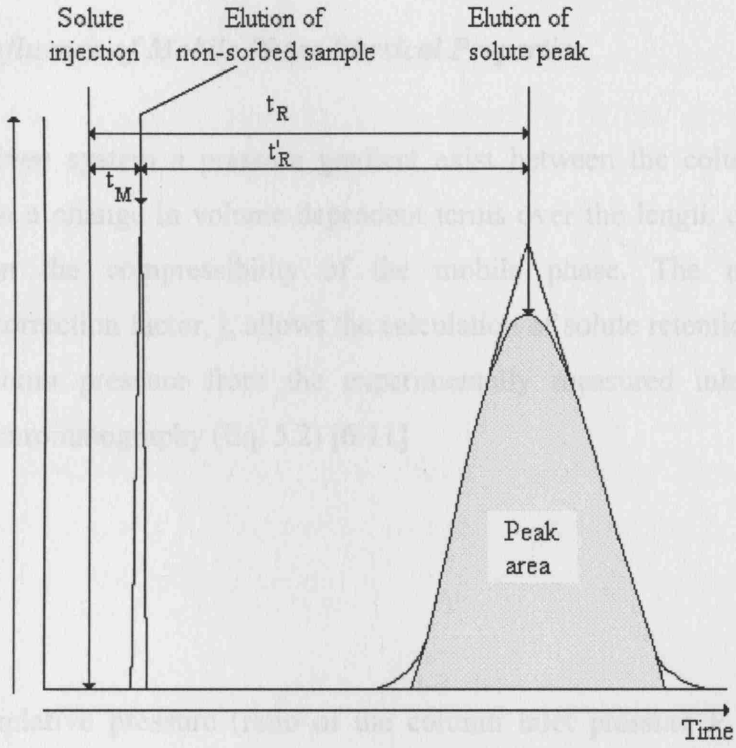


Figure 5.2. Example of a Chromatogram.

5.1. RETENTION IN GAS-LIQUID CHROMATOGRAPHY

The position of the peak in a chromatogram is characterized by its retention time (t_R) or retention volume (V_R). Retention volumes are fundamentally more correct than time but require further experimental information for their determination. The retention time is made up of two components, the time that the solute spends in the mobile phase and the time it spends in the stationary phase [3], Figure 5.2.

The time required by the mobile phase entering the column to reach the detector is called the column hold-up time, t_M, which in volume terms is equivalent to the volume of streaming mobile phase contained in the column.

The time the solutes spends in the stationary phase is called the adjusted retention time, t_R', (or adjusted retention volume, V_R') and is calculated by difference from the retention time (volume) and the column hold-up time (volume) [4]. Since for

convenience the retention time of a substance is determined from the moment of injection as time zero, the simple relationship (Eq. 5.1) combining the independent contributions to the observed retention time can be deduced.

$$t_R = t_M + t_R' \tag{5.1}$$

5.1.1 Influence of Mobile Phase Physical Properties

In a pressure-driven system a pressure gradient exist between the column inlet and outlet resulting in a change in volume-dependent terms over the length of the column that depends on the compressibility of the mobile phase. The mobile phase compressibility correction factor, j, allows the calculation of solute retention volumes at the average column pressure from the experimentally measured inlet and outlet pressures in gas chromatography (Eq. 5.2) [6-11]

$$j = \frac{3}{2}\left[\frac{(P^2-1)}{(P^3-1)}\right] \tag{5.2}$$

where P is the relative pressure (ratio of the column inlet pressure P_i to the outlet pressure P_0) and $P_i = P_G$ (column head pressure) + P_0 (ambient pressure).

If the measurement to calculate the carrier gas flow rate at the column temperature from the flow rate measured at the column outlet F_0 is made with a soap-film meter it is necessary to correct the flow rate for the difference between the dry gas (column) and water saturated gas (meter) measurements. F_c is the corrected carrier gas flow rate (Eq. 5.3)

$$F_c = F_0\left(\frac{T_c}{T_a}\right)\left[1-\left(\frac{P_w}{P_0}\right)\right] \tag{5.3}$$

where P_w is the vapour pressure of water at ambient temperature, T_a, and T_c is the column temperature.

The retention volume (V_R) is the volume of mobile phase entering the column between sample injection and the emergence of the peak maximum for the substance of

interest (Eq. 5.4). The corrected retention volume ($V_R°$) is the retention volume corrected for the compressibility of the carrier gas (Eq. 5.5)

$$V_R = t_R F_c \tag{5.4}$$

$$V_R° = jV_R \tag{5.5}$$

The adjusted retention volume (V_R') is the retention volume corresponding to the adjusted retention time and the net retention volume (V_N) is the adjusted retention volume corrected for the compressibility of the carrier gas.

$$V_R' = t_R' F_c \tag{5.6}$$

$$V_N = jV_R' \tag{5.7}$$

The specific retention volume ($V_g°$) is the net retention volume per gram of stationary phase (either liquid phase or solid phase) at the column temperature (Eq. 5.8). This is commonly used in the determination of physicochemical properties by gas chromatography.

$$V_g° = \frac{V_N}{W_L} \tag{5.8}$$

W_L is the weight of liquid phase or solid phase. The specific retention volume corrected to 0° C is given by Eq. (5.9)

$$V_g = V_g° \left(\frac{273.2}{T_c} \right) \tag{5.9}$$

For the most exact work, it may be necessary to allow for non-ideal behaviour of the gas phase by applying a correction for solute-gas phase interactions [5,12]. For carrier gases that are insoluble in the stationary phase and at moderate column inlet pressures (Eq. 5.10) is a reasonable approximation

$$\ln V_N = \ln V_N(0) + 0.75\left[\frac{(2B_{12} - V_1)}{RT_c}\right]\left[\frac{(P^4 - 1)}{(P^3 - 1)}\right]P_0 \qquad (5.10)$$

where $V_N(0)$ is the net retention volume at zero column pressure drop, B_{12} is the second interaction virial coefficient of the solute with the carrier gas, V_1 is the solute molar volume at infinite dilution in the stationary phase (commonly replaced by the bulk molar volume), R is the universal gas constant, P is the relative pressure and P_0 is the pressure at the column outlet. Under normal operating conditions, errors due to assuming ideality of the gas phase for simple carrier gas like hydrogen, helium and nitrogen are small, however, they increase with high solute concentrations, large column pressure drops, and low temperatures.

5.1.2 Property Estimations

Gas chromatography is widely used to determine solution and adsorption thermodynamic properties [4,5,12-17]. The free energy, enthalpy, and entropy of mixing or solution, and infinite dilution solute activity coefficients can be determined from retention measurements made at infinite dilution (Henry's law region) in which the value of the Raoult's activity coefficient, γ_1^∞, (also the gas-liquid distribution constant) can be assumed to have a constant value.

The activity coefficient and the specific retention volume are related by Eq. (5.11)

$$V_g^\circ = \frac{R}{M_2 \gamma_1^\infty P_1^\circ} \qquad (5.11)$$

where M_2 is the molecular weight of the solvent, γ_1^∞ the solute activity coefficient at infinite dilution, and P_1° the saturation vapor pressure of the pure solute at the given temperature.

The gas-liquid distribution constant (K_L), moles of solute per unit volume of liquid/moles of solute per unit volume of gas phase, is evaluated from the specific retention volume using the relationship (Eq. 5.12)

$$V_g^\circ = \frac{K_L}{\rho_c} \tag{5.12}$$

where ρ_c is the liquid phase density at the column temperature [17].

Also the partial molar Gibbs free energy of solution for a solute at infinite dilution (ΔG°) in the stationary phase can be obtained directly from the gas-liquid distribution constant by Eq. (5.13)

$$\Delta G^\circ = -RT_c \ln K_L \tag{5.13}$$

From the slope of a plot of log V_g° against the reciprocal of the column temperature over a narrow temperature range, 10-30 K, the enthalpy of solution is obtained. The entropy for the same process is obtained from a single value of the specific retention volume and the value of the enthalpy of solution calculated as just described [18-20].

5.2 RETENTION INDEX SYSTEMS AND THEIR APPLICATIONS

Since the early days of gas chromatography an enormous effort has been devoted to standardizing methods used to determine retention data to make possible the wider use of collected published results for identification purposes. This problem was solved by Kovats who proposed the retention index scale, in which the retention behaviour of any substance was expressed with respect to the retention properties of a series of closely related standard substance [21]. The standard substances were the n-alkanes, each one of which was assigned an index value of 100 times its carbon number (100 for methane, 200 for ethane, etc,) at any temperature and on any phase.

The retention index is represented by the symbol I and is calculated under isothermal conditions using Eq. (5.14)

$$I = 100z + 100\frac{[\log t_R'(x) - \log t_R'(z)]}{[\log t_R'(z+1) - \log t_R'(z)]} \tag{5.14}$$

where t_R' is the adjusted retention time, z the carbon number of the n-alkanes eluting immediately before the substance of interest denoted by x, and z + 1 the carbon number

of the n-alkane eluting immediately after substance x. Thus the retention index of a substance is expressed on a uniform scale with increased precision due to the use of two closely eluting and bracketing standards for the experimental measurement. For standardization of reporting it is recommended that the retention index be indicated by I with the stationary phase identified as a superscript, the sample as a subscript, followed by the temperature of the measurement in parenthesis.

Eq. (5.14) is the most common expression for the retention index, although the same values are obtained by replacing the adjusted retention times by the specific retention volumes, gas-liquid partition coefficients, or the Gibbs free energies of solution when the retention mechanism is gas-liquid partitioning [21,22].

5.2.1 Retention index standards

The standard substances were initially the n-alkanes, as mentioned. However, the retention index concept is not restricted to use of n alkanes as index standard and any homologous series that shows a linear relationship between the logarithm of the adjusted retention time and carbon number can be used [23]. Standards of intermediate polarity such as 2-alkanones, alkyl halides, alkyl acetates and alkanoic acid methyl esters are some examples. Other examples are index standards which are chosen by compatibility with a selective detector such as n-alkyltrichloroacetates, n-bromoalkanes, 1-nitroalkanes and dialkylsulfides.

5.2.2 Temperature dependence of Kovats indices

In current gas chromatography it is possible to achieve high precision of retention index measurement, the repeatability of the experimentally obtained retention values could be ± 0.1 index unit or even better. However, the interlaboratory reproducibility of retention indices is worse [24]. Low interlaboratory reproducibility of the retention values restricts widespread use of data banks of retention indices in chromatographic practice. The major problem is associated with the determination and/or calculation of gas hold-up (dead) time. The other very important reason for the poor interlaboratory reproducibility seems to lie in absorption of the compounds during analysis. A limitation to the larger scale use of retention indices is that the values of studied compounds deal with are determined experimentally for a given column and only one

temperature.

When an isothermal index value is required at a temperature other than the temperature it was measured at, it can be obtained by interpolation using several empirical relationships like the hyperbolic function (Antoine-type equation)

$$I(T) = A + \frac{B}{(T + C)} \tag{5.15}$$

where I(T) is the retention index at any temperature T(K) and A, B and C are experimentally derived constants [25,26]. The curve can nevertheless have a significant linear portion, particularly for substances of low polarity on non-polar stationary phases.

For mixtures of a wide boiling point range the determination of retention indices under isothermal conditions would be time consuming and unnecessarily restrictive. Under temperature program conditions, there is an approximately linear relationship between the n-alkane elution temperature and their carbon number [21,27,28]:

$$I = 100z + 100\frac{[T_R(x) - T_R(z)]}{[T_R(z+1) - T_R(z)]} \tag{5.16}$$

where T_R is the elution temperature (K) and the subscripts x, z, z + 1 are identified in the Eq. (5.12).

Temperature program retention indices are more sensitive to the chromatographic conditions than isothermal indices and are generally of lower precision. Temperature program indices are influenced by the characteristics (time and temperature) of any isothermal period before the start of the program, the temperature program rate (faster program rates lead to higher elution temperatures), and differences in column dimensions [29].

In addition to the individual retention index values, the retention index difference is of great significance because of their physicochemical meaning

$$\Delta I(T) = I_s^p(T) - I_s^{Np}(T) \tag{5.17}$$

where ΔI is the difference between the retention indices of a single substance(s)

measured on two different stationary phases (p and Np) at an identical, isothermal column temperature (T), p is a polar stationary phase and Np a non-polar stationary phase, e.g., squalane.

This retention index difference has significant practical importance. The ΔI values can be approximately used for the calculation of individual relationship, to follow their variation, for the characterization of the polarity of stationary phases, to predict the values of retention data, etc [21].

5.2.3 Unified Retention Index

The actual relationship between retention index (I) and column temperature (T) can be described by Eq. (5.15). As mentioned, this equation corresponds to a hyperbola; however, it has been demonstrated [21,31] that in the temperature interval used in practice the I vs T plot has a linear section and therefore Eq (5.18) can be written

$$I \approx aT + b \tag{5.18}$$

where a and b are constants. The linear relationship generally covers a larger temperature range with a non-polar and polar phase.

Vernon and Suratman [32] demonstrated that for the temperature ranges used in practice, the linear relationship between I and T may be written as

$$I = I_0 + mT \tag{5.19}$$

where I_0 is the theoretical retention index at 0°C and T the working temperature in °C.

Due to the approximate linear relationship described in Eq. (5.12) and Eq. (5.19), one can express the temperature dependency of the retention index by giving the increment for a given temperature range, usually 10°C, and then use this value to calculate the retention index at selected temperatures [24]. In recent years, the stated conclusion has been applied by Dimov [33] and Skrbic [34] in the form of

$$UI_T = UI_0 + \left(\frac{dUI}{dT}\right)T \tag{5.20}$$

where UI_T is the unified retention index at temperature T (°C), UI_0 is the value of UI_T at 0°C, dUI/dT is the index increment with the analysis temperature (usually given as dUI/10°C).

The statistical treatment using simple regression analysis of the experimental data allows computation of a unified retention index (UI_T) by Eq. (5.20). The values of unified retention index (UI_T) obtained and its temperature increment were considered as reliable if the data included in the regression matrix were from two authors and at three temperatures at least, and no more than 33% of all data were excluded [24].

5.2.4 Connection between the Physicochemical Quantities and Retention Indices

In practice, the following relationship exists between the specific retention volume and the retention indices

$$\log V_g(s)_T = \left[\frac{I_s^{st.ph.}(T) - 100z}{100}\right] b_T^{st.ph.} + \log V_g(z) \tag{5.21}$$

where b is the slope of the n-alkane curve

$$b = \log t_R'(z+1) - \log t_R'(z) \tag{5.22}$$

and z the carbon number of the n-alkane. By the n-alkane curve is meant the plot of the log t_R'(z) versus carbon number (z) relationship for n-alkanes

$$\log t_R'(z) = bz + a \tag{5.23}$$

where a is the intercept on the ordinate. This plot is generally linear except for the first few members of the homologous series.

In order to eliminate the error in the gas hold-up time determination or calculation the following relationship has been proposed for the determination of the b value [30], if $z \geq 7$

$$b = \log\left[\frac{t_R(z+2) - t_R(z+1)}{t_R(z+1) - t_R(z)}\right] \tag{5.24}$$

The temperature dependence of b can be approximately described with the following equation

$$b_T^{st.ph.} = \frac{D}{T} + E \tag{5.25}$$

where D and E are constant and T is the column temperature (K).

A useful relationship is available if the following data of the gas chromatographic system are known:

$$\log V_g(s_2) = \left[\frac{I_{s_2}^{st.ph}(T) - I_{s_1}^{st.ph}(T)}{100}\right] b_T^{st.ph.} + \log V_g(s_1) \tag{5.26}$$

A well known relationship exists between the partition coefficient and the specific retention volume [17], as mentioned

$$K_L = V_g^{st.ph.} \rho_c \tag{5.27}$$

where ρ_c, is identified in Eq. (5.12), and $V_g^{st.ph.}$ is equal to $V_g°$.

5.3. QUALITATIVE AND QUANTITATIVE ANALYSIS

5.3.1 Qualitative Analysis

Chromatography is, first and foremost, a separation method, and, although peaks may be assigned a tentative identification by means of retention data, it is often essential to use other independent method of identification [35]. That is, each chromatographic peak can be directed into a mass spectrometer or Fourier transform infrared spectrophotometer to record a spectrum as the substance is eluted from the column. The spectrum can be identified by comparison with a library of spectra stored in a computer.

Hyphenated analytical methods are often used because multidimensional information can be obtained from a single analysis. gas chromatography is the most common technique for separation of volatile and semi-volatile mixtures. It is well accepted that when gas chromatography is coupled with spectral detection methods, such as mass spectrometry (MS) or Fourier transform infrared (FT-IR) spectrometry, that the resulting combination is a powerful tool for the analysis of complex mixtures. In particular, multispectral analysis systems are even more definitive than those employing a single spectral detector [36].

5.3.2 Quantitative Analysis

For quantitative analysis is necessary to establish a relationship between the magnitude of the detector signal and the sample amount. The detector signal is measured by the peak area or height from the chromatogram. A number of manual methods are available for calculating peak areas, Table 5.1 [3]. No single method is perfect and common problems include the difficulty of accurately defining peaks boundaries, variable precision between analysts and the need for a finite time to make each measurement. A major disadvantage of manual measurements is the necessity that all the peaks of interest must be completely contained on the chart paper.

Table 5.1
Comparison of manual methods for determining peak areas
H = peak height, H' = peak height by construction (see Fig. 5.2) and w_h = peak width at half height, w_b at base, $w_{0.15}$ at 0.15H, $w_{0.25}$ at 0.25H and $w_{0.75}$ at 0.75H

Method	Calculation	True area (%)	Precision (%)
Gaussian peak	Hw_h	93.9	2.5
Triangulation	$H'w_b/2$	96.8	4.0
Condol-Bosch	$H(w_{0.15} + w_{0.75})/2$	100.4	2.0
EMG	$0.73Hw_{0.25}$	100	2.0
Planimetry		100	4.0
Cut and weigh		100	1.7

Adapted from Ref. 3

Dedicated electronic integrators and personal computers with appropriate software for integration are routinely used for recording chromatograms and this has overcome the

problems above. However, it is important to understand the limitations of electronic integration [3].

Four techniques are commonly used to convert peak height or area information into relative composition data for the sample. These are normalisation method, the external standard method, the internal standard method and the method of standard additions [36].

5.4. CHARACTERIZATION OF A STATIONARY PHASE

5.4.1 Introduction

Separations are achieved in gas chromatography by distribution of a solute between an immobile solid or liquid stationary phase and a gas phase that percolates over stationary phase. The distribution constant is determined by the column temperature and the extent of intermolecular interactions between the solute and the stationary phase. The mobile phase is responsible for transport through the column but otherwise does not participate in the retention mechanism [3].

Useful separations require kinetically optimised conditions achieved through column design, and the optimisation of selectivity, which is achieved by either temperature variation or by using different stationary phases. The former is easily achieved by external means, but is rather limited in scope, since large changes in temperature result in either excessive or insufficient retention to provide practically useful separations, and small changes in temperature, in insufficient variation in the relative strengths of intermolecular interactions to provide wide changes in selectivity. Temperature remains an important and useful optimisation parameter in gas chromatography, but effective selectivity optimisation requires that a wide range of stationary phases distinguished by their capacity for varied intermolecular interactions be employed [37]. For this purpose it is important to have a reliable, systematic approach to stationary phase classification so that the selectivity differences between stationary phases can be identified.

5.4.2 Rohrschneider's and McReynolds' phase constants

The selectivity of a stationary phase is defined as its relative capacity for specific

intermolecular interactions, such as dispersion, induction, orientation and complexation (including hydrogen bond formation). Unlike "polarity" it should be possible to provide meaningful experimentally-derived scales for selectivity. The system of phase constants introduced by Rohrschneider [38,39] and later modified by McReynolds [40] was the first widely adopted approach for the systematic organization of stationary phases based on their selectivity for specific solute interactions

The founding principle of Rohrschnieder's method was that the intermolecular forces were assumed to be additive, and that their individual contributions to retention can be evaluated from the differences in retention index values for a series of selected solutes on the stationary phases to be characterized and on squalane, used as a non polar reference phase, at a fixed temperature [38,39]. Rohrschnieder assumed that the retention index of a substance on a non-polar phase, such as squalane, was determined solely by dispersion forces, and that any difference in that retention index values for a polar phase and a non-polar phase was due to polar interactions and could be expressed by Eq. (5.28)

$$I^{P}_{PH} = I^{P}_{SQ} + \Delta I \qquad (5.28)$$

where I^{P}_{PH} is the retention index for a selected solute P on the stationary phase to be characterized, PH; I^{P}_{SQ} the retention index for the same solute on the non-polar reference phase squalane, SQ; and ΔI is the retention index difference equivalent to the contribution from polar interactions.

These retention index differences representing polar interactions were represented by Rohrschnieder as a series of terms composed of a solute specific contribution (a, ..., e) and a stationary phase characteristic term (X', ..., S') allowing the retention index difference to be written as

$$\Delta I = aX'+bY'+cZ'+dU'+eS' \qquad (5.29)$$

Five solutes, Table 5.2, were selected to characterize the principal intermolecular interactions responsible for retention, i.e., solutes for which all a-e parameters are nil but one.

Table 5.2
Characteristic solute interactions associated with the first five McReynolds' probes (Rohrschneider's probes in parentheses)

Symbol	Solute	Interaction measured
X'	Benzene	Primarily dispersion with weak hydrogen-bond base properties. Polarizable in induction interactions.
Y'	Butanol (ethanol)	Orientation with both hydrogen-bond acid and base capabilities.
Z'	2-Pentanone (2-butanone)	Orientation with some weak hydrogen-bond and base capabilities. No hydrogen-bond acidity.
U'	Nitropropane (Nitromethane)	Orientation with some weak hydrogen-bond basicity. No hydrogen-bond acidity.
S'	Pyridine	Strong hydrogen-bond base with moderate orientation capability. No hydrogen-bond acidity.

Adapted from Ref. 37

Determination of the stationary phase characteristic constants was performed by measuring ΔI for each solute in turn and assigning a value of 100 for the solute specific constant associated with that solute. Thus, by determining the retention index for five selected solutes on the stationary phase to be characterized, and on squalane, at a fixed reference temperature of 100°C, it was possible to assign numerical values for the five stationary phase characteristic constants. McReynolds modified Rohrschneider's approach by increasing the number of solutes used in Eq. (5.29) from five to ten, replacing three of the Rohrschneider's solutes (ethanol, nitromethane, 2-butanone) by less volatile homologues (n-butanol, nitropropane and 2-pentanone) to provide easier access to retention index values on some phases where Rohrschneider's solutes possessed low retention, proposed 120°C instead of 100°C as the standard reference temperature, and used ΔI instead of $\Delta I/100$ for the calculation of the characteristic phase constants. It is McReynolds' characteristics phase constants that are usually found in the contemporary scientific and commercial literature, see Table 5.3 for some examples.

It is now realized for a number of fundamental and practical reasons, reviewed elsewhere [41,42,43,44], that Rohrschneider's approach is unsound for stationary phases characterization by their selectivity for specific interactions.

5.4.3 Solvation parameter model

Modern approaches to stationary phase characterization by selectivity ranking are based on the cavity model of solvation [45,46]. This model, as mentioned, assumes that the

transfer of a solute from the gas phase to solution in the stationary phase involves three steps (i) the endoergic creation of a cavity in the solvent, (ii) reorganization of the solvent molecules around the cavity, and (iii) incorporation of the solute in the cavity with setting up of various exoergic solute-solvent interactions. The individual free energy terms involved in the transfer are assumed to be additive. Each of those interactions require of a relevant solute parameter or descriptor. After considerable preliminary work [47], the following solute descriptors were selected: R_2, π_2^H, α_2^H, β_2^H, L^{16}, and V_x, actually represented by **E, S, A, B, L** and **V**, respectively, to simplify representation of the model [51,52]. They were described above in section 2.2 of this work. Briefly,

- **E** is an excess molar refraction that can be calculated from refractive index or can be rather easily be estimated [47],
- **S** is the solute dipolarity-polarizability obtained to date from gas-liquid chromatography (GLC) of solutes on polar stationary phases [48],
- **A** and **B** are the solute hydrogen-bond acidity and solute hydrogen-bond basicity (where appropriate $\Sigma\alpha_2^H$ and $\Sigma\beta_2^H$ must be used) [45],
- **L** is the solute gas-hexadecane partition coefficient at 298 K [49], and
- **V** is McGowan's characteristic volume [50].

Table 5.3
Stationary Phases and McReynold's Values

Phase	X'	Y'	Z'	U'	S'	Average
Apiezon L	32	22	15	32	42	28.6
OV-3	44	86	81	124	88	84.6
Squalene	152	341	238	329	344	280.8
Carbowax 20	322	536	368	572	510	461.6
DEGS	496	746	590	837	835	700.8

Adapted from Ref. 56

Two general linear solvation energy relationships (LSERs) can be constructed from these parameters [48],

$$SP = c + e \cdot \mathbf{E} + s \cdot \mathbf{S} + a \cdot \mathbf{A} + b \cdot \mathbf{B} + v \cdot \mathbf{V} \quad (5.30)$$

$$SP = c + e \cdot \mathbf{E} + s \cdot \mathbf{S} + a \cdot \mathbf{A} + b \cdot \mathbf{B} + l \cdot \mathbf{L} \quad (5.31)$$

In Eq. (5.31), the dependent variable SP refers to gas chromatographic data for a series of solutes on a stationary phase (or phases), for example log V_g or log t_{rel} or I. The system constants in the model (Eq. 5.31) convey all information of the ability of the stationary phases to participate in solute-solvent intermolecular interactions. The constants are found by the method of multiple linear regression analysis.

- *e* constant refers to the ability of the stationary phase to interact with solute n- or π–electron pairs,
- *s* constant establishes the ability of the stationary phase to take part in dipole-type interactions,
- *a* constant is a measure of stationary phase hydrogen-bond basicity (because a basic phase will interact with an acidic solute,
- *b* constant is a measure of stationary phase hydrogen-bond acidity, and
- *l* constant incorporates contributions from stationary phase cavity formation and solute-solvent dispersion interactions.

The coefficients in the solvation parameter equation are not just fitting constants, but must obey general chemical principles. It should be noted that the regression equation remains the same except for the c-constant, no matter whether the dependent variable is log K_L (log L), or log V_g or even log t_{rel}. But the regression coefficients do not remain the same if I is used.

Since the system constants change markedly with temperature [37], comparison of phases is best made at a common temperature. For instance, in general polar interaction are expected to decrease with increasing temperature, and cavity formation should be easier at higher temperatures, but for individual stationary phases the magnitude of these temperature-induced changes will not be the same. Out of practical necessity the solute descriptors are considered temperature invariant (otherwise a new set of descriptor would be needed for each temperature used).

The system constants for some common packed columns are summarized in Table 5.4 [37,46,53].

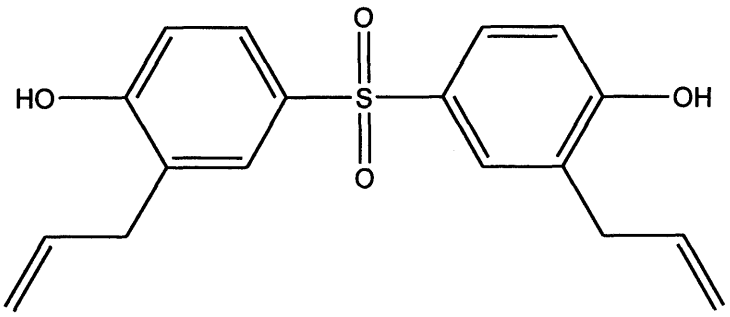
Table 5.4
System constant derived from the solvation parameter model for packed column stationary phases (120-122°C)

Stationary phases	System constant					
	e	*s*	*a*	*b*	*l*	*c*
Squalane (SQ)	0.129	0.011	0	0	0.583	-0.222
Apolane 87	0.170	0	0	0	0.549	-0.221
Carbowax 20M CW20M	0.317	1.256	1.883	0	0.447	-0.560
1,2,3-Tris(2cyanoethoxypropane) TCEP	0.116	2.088	2.095	0.261	0.370	-0.744
Poly(diethylene glycol succinate) DEGS	0.230	1.572	2.105	0.171	0.407	-0.650
Poly(dimethylsiloxane) SE-30	0.024	0.190	0.125	0	0.498	-0.194
Poly(dimethylmethylphenylsiloxane) OV-3	0.033	0.328	0.152	0	0.503	-0.181
Bis(3-allyl-4-hydroxyphenyl)sulfone H10	-0.051	1.323	1.266	1.457	0.418	-0.568

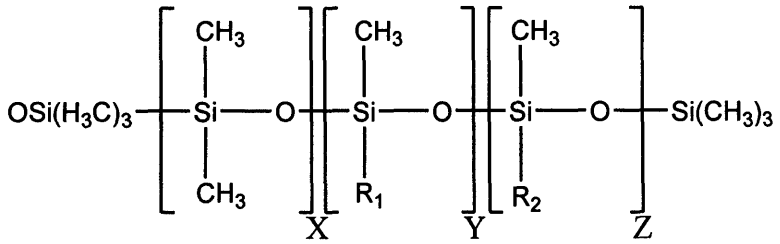
It can be seen from Table 5.4 that squalane, SE-30 and OV-3 are all rather non-polar with small or negligible hydrogen-bond basicity. Carbowax and TCEP, for example, are both polar/polarizable and of considerable hydrogen-bond basicity. It is only recently that suitable stationary phases with hydrogen-bond acid properties have been described. The phenolic stationary phase of Abraham et al. [54], H10, is an example of these stationary phases, it is a reasonably strong hydrogen-bond acid, but also has considerable hydrogen-bond basicity. This is in line with the presence of the strongly basic sulfone functionality as shown in Figure 5.3. Outstanding among acid phases is PSF6 which has a wide temperature operating range and is a strong hydrogen-bond acid with zero hydrogen-bond basicity. Because hydrogen-bond acid stationary phase are new to gas chromatography and none are commercially available, there is little experience in their use for selectivity optimisation.

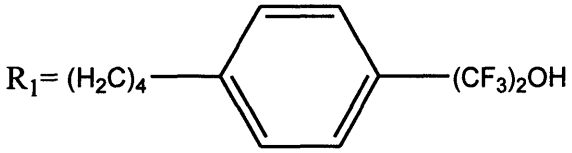
Chemometric classification procedures allows the evaluation of the extent of the separation space that can be employed for separations in gas chromatography based on the range of properties of stationary phases that have been characterized by the solvation parameter model. Principal component analysis and hierarchical clustering methods provide useful approaches for stationary phases classification [37,53,55].

A



B





$R_2 = (CH_2)_7CH_3$

X = 70% Y = 26% Z = 4%

Figure 5.3. Structures of two hydrogen-bond stationary phases. A, Bis3-allyl-4-hydroxyphenyl)sulfone (H-10) and B, Poly(oxy{methyl[4-(2hydroxy-1,1,1,3,3,3-hexafluoropropyl-2-yl)phenyl]butyl}silylene)-co-oxy(dimethylsilylene) containing 4 mol % octylmethysiloxane and 70 % dimethylsiloxane (PSF6).
Adapted from Ref. 3

5.5. REFERENCES

1. L. S. Ettre, Pure & Appl. Chem. 65 (1993) 819
2. L. S. Ettre, Chromatographia 38 (1994) 521
3. C. F. Poole, (2003) *The Essence of Chromatography*, Elsevier Science B.V, Amsterdam
4. R. L. Grob, E. F. Barry (eds.), (2004) *Modern Practice of Gas Chromatography*, 4th ed., Wiley, New York
5. J. A. G. Dominguez, J. C. Diez-Masa, V. A. Davankov, Pure Appl. Chem. 73 (2001) 969

6. L. S. Ettre, J. V. Hinshaw, Chromatographia 43 (1996) 159
7. L. M. Blumberg, Chromatographia 44 (1997) 326
8. J. F. Parcher, Chromatographia 47 (1997) 326
9. V. A. Davankov, L. A. Onuchak, S. Yu. Kudryashov, Yu. I. Arutyunov, Chromatographia 49 (1999) 449
10. G. Foti, E. Kovats, J. High Resolut. Chromatogr. 23 (2000) 119
11. J. R. Conder, C. L. Young, (1979) *Physicochemical measurement by gas chromatography*, Chichester. Wiley, New York
12. R. J. Laub, P. L. Pecsok, (1978) *Physicochemical Applications of gas chromatography*, Wiley, New York
13. N. A. Katsanos, F. Roubani-Kalantzopoulou, Adv. Chromatogr. 40 (2000) 231
14. N. A. Katsanos, R. Thede, F. Roubani-Kalantzopoulou, J. Chromatogr. A 795 (1998) 133
15. D.C. Locke, Adv. Chromatogr. 14 (1976) 87
16. K. Kojima, S. Zhang, T. Hiaki, Fluid Phase Equilibria 131 (1997) 145
17. F. R. Gonzalez, J. Chromatogr. A 942 (2002) 211
18. K. G. Furton, C. F. Poole, J. Chromatogr. 399 (1987) 47
19. E. F. Meyer, J. Chem. Edu. 50 (1973) 191
20. R. C. Castells, J. Chromatogr. 350 (1985) 339
21. G. Tarjan, Sz. Nyiredy, M. Gyor, E. R. Lombosi, T. S. Lombosi, M. V. Budahegyi, S. Y Meszaros, J. M. Tackas, J. Chromatogr. 472 (1989) 1
22. S. K. Poole, B. R. Kersten, C. F. Poole, J. Chromatogr. 471 (1989) 91
23. G. Castello, J. Chromatogr. A 842 (1999) 51
24. B. D. Skrbic, J. Chromatogr. A 764 (1997) 257
25. E. Tudor, J. Chromatogr. A 858 (1999) 65
26. J. P. Chen, X. M. Liang, Q. Zhang, L. F. Zhang, Chromatographia 53 (2001) 548
27. F. R. Gonzalez, A. M. Nardillo, J. Chromatogr. A 842 (1999) 29
28. F. R. Gonzalez, L. G. Gagliardi, J. Chromatogr. A 879 (2000) 157
29. C. Bicchi, A. Binello, A. D'Amato, P. Rubiolo, J. Chromatogr. Sci. 37 (1999) 288
30. T. Toth, J. Takacs, Magy. Kem. Foly 77 (1971) 252
31. P. Chovin, J. Lebbe, J. Gas Chromatography 4 (1966) 37
32. I. Vernon, J. B. Suratman, Chormatographia 12 (1983) 600
33. N. Dimov, J. Chromatogr. 347 (1985) 366
34. B. D. Skrbic, J. Dj. Cvejanov, Lj. S. Pavic-Suzuki, Chromatographia 42 (1996) 660

35. A. J. Handley, E. R. Adlard, (1995) *Gas Chromatography. Techniques and Applications*, Shelfied Academic Press, PLACE
36. Y. T. A. Sasaki , C. L. Wilkins, J. Chromatogr. A 842 (1999) 341
37. M. H. Abraham, C. F.Poole, S. K. Poole, J. Chromatogr. A 842 (1999) 79
38. L. Rohrschneider, J. Chromatogr. Sci. 11 (1973) 160
39. L. Rohrschneider, Chromatographia 38 (1994) 679
40. W. O. Mc Reynolds, J. Chromatogr. Sci. 8 (1970) 685
41. V. G. Berezkin, A. A. Korolev, I. V. Malyukova, J. High Resolut. Chromatogr. 20 (1997) 333
42. C. F. Poole, S. K. Poole, Chem. Rev. 89 (1989) 377
43. B. R. Kersten, C. F. Poole, J. Chromatogr. 452 (1988) 191
44. C. F. Poole, T. O. Kollie, S. K. Poole, Chromatographia 34 (1992) 281
45. M. H. Abraham, Chem. Sov. Rev. 22 (1993) 73
46. C. F. Poole, S. K. Poole, J. Chromatogr. A 965 (2002) 263
47. M. H. Abraham, G. S. Whiting, R. M. Doherty, W. J. Shuely, J. Chromatogr. 587 (1991) 213
48. M. H. Abraham, G. S. Whiting, R. M. Doherty, W. J. Shuely, J. Chem Soc., Perkin Trans. 2 (1990) 1451
49. M. H. Abraham, P. L Grallier, R. A. McGill, J. Chem Soc., Perkin Trans. 2 (1987) 797
50. M. H. Abraham, J. C. McGowan, Chromatographia 23 (1987) 243
51. M. H. Abraham, C. M. Du, J. A. Platts, J. Org. Chem. 65 (2000) 7114
52. M. H. Abraham, J. A. Platts, J. Org. Chem. 66 (2001) 3484
53. S. K. Poole, C. F. Poole, J. Chromatogr. 697 (1995) 415
54. M. H. Abraham, J. Andonian-Haftvan, I. Hamerton, C. F. Poole, T. O. Kollie, J. Chromatogr. 646 (1993) 351
55. S. D. Martin, C. F. Poole , M. H. Abraham, J. Chromatogr A 805 (1998) 217–235
56. The Retention Index System in Gas Chromatography: McReynolds Constants, Available: http://www.sigma-aldrich.com (Accessed: March 3, 2006)

6.0. INTRODUCTION

The first systematic study of distribution between two immiscible liquids which led to a theory with predictive capabilities was carried out by Berthelot and Jungfleish [1]. The first appreciation was the basic fact that the ratio of the concentrations of solute distributed between two immiscible solvents was a constant and did not depend on the relative volumes of solutions used. It was concluded from these early observations that there was a small variation in partition coefficient with temperature.

In 1891, Nernst stressed the fact that the partition coefficient would be constant only if a single molecular species was being considered as partitioned between the two phases [2].

During the two decades bracketing the turn of the century, while the partition coefficient was being studied by physical chemists as an end in itself, pharmacologists became quite interested in the partition coefficient through the work of Meyer [3] and Overton [4] who showed that the relative narcotic activities of drugs towards the tadpole often paralleled their oil/water partition coefficients. But it was not until a little later that the partition coefficient began to be used as an extra-thermodynamic reference parameter for "hydrophobic bonding" in biochemical and pharmacological systems.

Partition coefficients are used extensively in medicinal chemistry, drug design, ecotoxicology and environmental chemistry. Nowadays, the partition coefficient is recognized by governmental and international agencies (U.S. Environmental Protection Agency, OECD) as a physical property of organic pollutants equal in importance to vapour pressure, water solubility and toxicity [5].

With the advance of modern techniques such as combinatorial chemistry, high throughput screening has become very important and therefore estimates of physical properties from structure by calculations that can be performed rapidly are of primary importance.

There is some time confusion over the definition and representation of the octanol-water partition coefficient. It is most often indicated as log P, log K, etc. frequently with a subscript o or w. Log P is generally used to indicate any partition

coefficient and log K any equilibrium constant.

6.1. THEORETICAL

6.1.1 Henry's Law

The most general approach to distribution phenomena is to treat the partition law as an extension of Henry's law. For a gas in equilibrium with its solution in some solvent

$$K = \frac{m}{p} \tag{6.1}$$

where m is the mass of gas dissolved per unit volume and p is the pressure at constant temperature. Since the concentration of molecules in the gaseous phase is proportional to pressure, p can be replaced by C_1 and the mass/unit volume of gas in solution designated by C_2. Eq. (6.1) can then be restated as

$$K = \frac{C_2}{C_1} \tag{6.2}$$

If C_1 and C_2 are in the same units, e.g. mol·dm^{-3}, the K is the unitless gas-to-water partition (or equilibrium) coefficient. The definition of the thermodynamic equilibrium constant follows from Eq. (6.2) as shown later.

In the most general terms, then, the concentrations of any singular molecular species in two phases which are in equilibrium with one another will bear a constant ratio to each other as long as its activity coefficients in the two phases are the same.

Many large interesting organic compounds deviate considerably from ideal behaviour in water and various solvents so that one is not always even reasonably sure of the exact nature of the molecular species undergoing partitioning.

6.1.2 Nonideal behavior of solutes

In many instances solute molecules can exist in different forms in the two phases. This problem could be illustrated with a compound studied in this work, fluorescein. The

chemical structures of the fluorescein protolytic forms are shown in Figure 6.1.

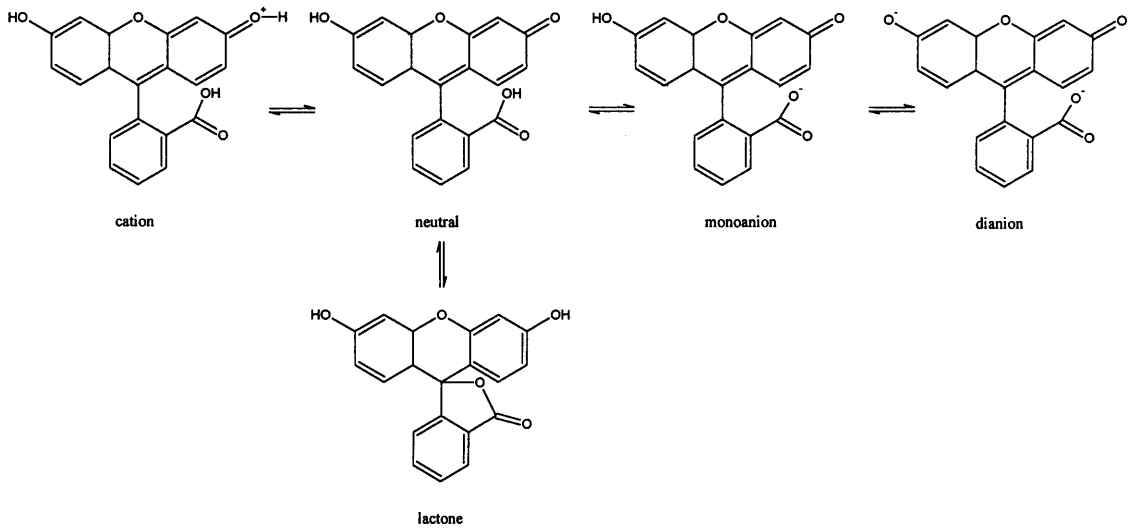


Figure 6.1. Ionization equilibria of fluorescein.

Fluorescein in aqueous solution occurs in cationic, neutral, anionic and dianionic forms, therefore, Henry's Law is not obeyed in this example. According to the Nernst law the true partition coefficient (log P) refers to the concentration ratio of unionized, monomer form of an organic compound in the two phases. For ionizable molecules the partition coefficient experimentally measured at a certain pH is an apparent value (log D). The relationships between log P and log D are well known for monoprotic acids and bases (Eqs. (6.3) and (6.4), respectively) [6].

$$\log P = \log D + \log(1 + 10^{pH-pK_a}) \tag{6.3}$$

$$\log P = \log D + \log(1 + 10^{pK_a-pH}) \tag{6.4}$$

If the ionised form partitions into the organic phase, more complicated relationships are required, as is also the case for di-protic and tri-protic substance [7].

Quite often log D values are measured at the standard 'physiological' pH of 7.4 and are not corrected to log P values. If it is log P that is required, the larger the difference between pH and the pK_a, the larger the correction factor. The partitioning of zwitterions is more complex. The relative concentrations of the anionic (+), the cationic (-), and the zwitterionic (+/-) forms will depend on the pH, but it is the electronically neutral (o) species, which co-exists with the zwitterions (+/-), that has the maximum lipophilicity.

6.1.3 Thermodynamics of Partitioning Systems

Consider two solvents that are wholly or partly immiscible and which form two liquid phases, 1 and 2, of one solvent saturated by the other. When a solute A is introduced in such a biphasic liquid system, it distributes between the two phases. The chemical potential, μ_{1A}, is expressed by

$$\mu_{1A} = \mu_{1A}^{\circ} + RT \ln a_{1A} \tag{6.5}$$

where μ_{1A}° is the standard chemical potential of A in liquid phase 1, T is the absolute temperature and R is the gas constant. The quantity a_{1A} is called the activity of constituent A. Evidently $\mu_{1A} = \mu_{1A}^{\circ}$ when a_{1A}=1. Similarly, in the other phase, the chemical potential,
μ_{2A}, is:

$$\mu_{2A} = \mu_{2A}^{\circ} + RT \ln a_{2A} \tag{6.6}$$

If the chemical potential is not identical in the two phases, mass transfer of A occurs, and the activity changes so that the chemical potential of A becomes equal in both phases, i.e. the equilibrium is reached. At equilibrium the chemical potentials of the solute in both phases are equal, $\mu_{1A} = \mu_{2A}$, and hence,

$$\mu_{2A}^{\circ} - \mu_{1A}^{\circ} = -RT \ln\left(\frac{a_{2A}}{a_{1A}}\right) \tag{6.7}$$

Let c_A be the concentration (in mol $\cdot$ L^{-1}) of the solute A. Then it is possible to write the activity of A as $a_A = f_A c_A$, where f_A, a proportionality constant, is called the activity coefficient of A and is a number that accounts for deviations from ideal behaviour and c_A is the molar concentration of A. In this way, Eq. (6.7) can be rewritten as

$$(\mu_{2A}^{\circ})^{c} - (\mu_{1A}^{\circ})^{c} = -RT \ln\left(\frac{f_{2A} c_{2A}}{f_{1A} c_{1A}}\right) \tag{6.8}$$

The quantity $(\mu^{o}_{2A})^{c} - (\mu^{o}_{1A})^{c}$ is the standard Gibbs energy change ΔG^{o} for the process, and the latter is related to the thermodynamic equilibrium constant , K^{o} or (in the present case) P^{o}, by

$$\Delta G^{o} = -RT \ln K^{o} = -RT \ln P^{o} \tag{6.9}$$

The activity of a solute in a very dilute solution (concentrations of $\sim\leq 10^{-4}$ M for a non-electrolyte may be considered infinitely dilute for this purpose) may be taken equal to its molar concentration, i.e. the activity coefficient is unity. In this case, Eq.(6.8) can be restated as

$$\Delta G^{o} = -RT\ln\left(\frac{c_{2A}}{c_{1A}}\right) \tag{6.10}$$

The above equations are quite general. For the specific system of partition between two phases, ΔG^{o} is often denoted as the standard free energy of transfer of solute A from solvent 1 (saturated with solvent 2) to solvent 2 (saturated with solvent 1).

Assuming the standard molar enthalpy is constant over a limited temperature range, it follows that

$$\frac{\partial \ln P^{c}}{\partial T} = \frac{\Delta \overline{H}^{oc}}{RT^{2}} \tag{6.11}$$

where $\Delta \overline{H}^{oc}$ is the standard enthalpy of transfer between the two solvents. The plot of ln P versus $1/T^{2}$ (classical Van't Hoff plots) should produce a straight line with slope $\Delta \overline{H}^{o}/R$. Unfortunately, this is not that simple, particularly because the mutual solubility of the solvents is also temperature dependent. So it is necessary to consider the temperature in which the values of $\Delta \overline{H}^{o}$ are obtained, because for one solvent a change in the temperature means a change in the system. As a general rule, it is possible to consider that the effect of temperature on the P value is not great if the temperature change is not dramatic [5,8].

The entropy of transfer can of course, be calculated from Eq. (6.12)

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ \qquad (6.12)$$

Experience has shown that the partition coefficient at low solute concentration is usually not highly dependent on changes in miscibility of the two phases due to addition of the solute. Even with solvent pairs as miscible as isobutylalcohol-water, the effect is small with solutes at 0.1 M or less, and solvent pairs less miscible, such as chloroform-water, will easily tolerate 0.1 M solute without appreciable miscibility changes [6].

The results given by Leo et al. [6] of varying degrees of accuracy for a variety of solutes in different solvent systems show that the effect of temperature is usually of the order of 0.01 log unit/deg and may be either positive or negative. These Δ log P/deg values correspond to an enthalpy of partitioning of 28 kJ/mol.

Table 6.1
Temperature effect on log P

Solvent-water	Solute	Temp. (°C)	Δ log P/deg
Octanol	Hexanoic acid	4-22	$1.7\cdot10^{-3}$
	Octanoic acid	4-22	0.0
	n-Butylpyridinium bromide	4-22	$-1.0\cdot10^{-2}$
	n-Tetradecylpyridinium bromide	4-22	$-1.0\cdot10^{-2}$
Ethyl ether	Acetic acid	0-25	$-1.2\cdot10^{-3}$
Chloroform	Methylamine	18-32	$1.0\cdot10^{-2}$
Olive oil	Antipyrine	7-36.5	$1.2\cdot10^{-2}$
Benzene	o-Phenylenediamine	20-70	$3.4\cdot10^{-3}$
1-Hexanol	Succinic acid	20-60	$-0.5\cdot10^{-3}$
Heptane	p-Chloroaniline	15-35	$5.5\cdot10^{-3}$

Adapted from Ref. 7

Recently Lei et al. [9] presented a novel method of rapidly estimating the enthalpy of octanol-water phase transfer ΔH°_{tr}, that is, the temperature dependence of P, for organic compounds from retention time measurements on reversed phase HPLC-columns at variable temperatures. They observed that the values of $\Delta\overline{H}^\circ$ for all compounds which were analysed were between 13 and 32 kJ/mol confirming a relatively minor temperature dependence of the partitioning between water and octanol.

A similar equation to Eq. (6.10) for ΔG° can be written in terms of mole fraction and the development would be as follows

$$\mu_{2A}^{\circ} - \mu_{1A}^{\circ} = -RT\ln\left(\frac{a_{2A}}{a_{1A}}\right) = -RT\ln\left(\frac{f_{2A}x_{2A}}{f_{1A}x_{1A}}\right) \tag{6.13}$$

As $x \to 0$, that means at infinitely dilute liquid solution, f_{1A} and $f_{2A} \to 1$, thus this equation can be rewritten as

$$\mu_{2A}^{\circ} - \mu_{1A}^{\circ} = -RT\ln\left(\frac{x_{2A}}{x_{1A}}\right) \tag{6.14}$$

and in the same way in which Eq. (6.10) was reached, Eq. (6.15) can be obtained

$$(\Delta G^{\circ})^{x} = (\Delta\mu_{A}^{\circ})^{x} = -RT\ln K^{x} = -RT\ln P^{x} \tag{6.15}$$

and then Eq. (6.11) is rewritten as

$$\frac{\partial \ln P^{x}}{\partial T} = \frac{\Delta\bar{H}^{\circ x}}{RT^{2}} \tag{6.16}$$

Two different concentration scales have been used, so it is well to know the relation among them. From the definition of mole fraction, x, and molarity, c, we can convert from one to other. The relation that apply to dilute solutions ($x << 1$) are given here,

$$c = x \cdot \frac{1000\rho^{S}}{MW^{S}} \tag{6.17}$$

where ρ^{S} is the density of the solvent and MW^{S} is the molecular weight of the solvent. Then, for transfer from solvent 1 to solvent 2,

$$\log P^c = \log P^x + \log\left[\frac{1000\rho_2^S / MW_2^S}{1000\rho_1^S / MW_1^S}\right] \tag{6.18}$$

and then,

$$\frac{\Delta\bar{H}^{\circ c}}{RT^2} = \frac{\Delta\bar{H}^{\circ x}}{RT^2} + \frac{\partial \ln}{\partial T}\left(\frac{\rho_2^S / MW_2^S}{\rho_1^S / MW_1^S}\right) \tag{6.19}$$

Some solvent densities vary with temperature. As a result, $\Delta\bar{H}^{\circ x}$ will not be quite the same as $\Delta\bar{H}^{\circ c}$.

6.1.4 Energy Requirements for Phase Transfer

The relative roles of the various binding forces which determine the way a solute distributes itself between two phases has been examined elsewhere [10].

Hydrogen bonding is a main factor to consider in studying the energy requirements for phase transfer. This is of paramount importance in determining the character of both the solute and the organic solvent phase. Compounds such as alcohols, esters, and ketones have quite different properties than hydrocarbons. Moreover, as solvents, it is not simply the hydrogen bond character of the pure compound which must be considered. One must keep in mind that rather larger amounts of water are present in these oxygen-containing solvents when saturated during the partitioning process. For example, octanol dissolves in water only to the extent of 0.0045 M. However, the molar concentration of water in octanol is 2.30 [7]. Abraham, aware of the important role of the hydrogen-bonding in chemistry and biochemistry, and in order to understand and to interpret these effect quantitatively, separated out various possible solute-solvent interactions, and established numerical scales for solute properties such as hydrogen-bond acidity and hydrogen bond basicity [11].

Certain solvents such as alcohols and amines can act as both donors and acceptors in hydrogen bonding. This increases their versatility as solvents. For this reason Meyer and Hemmi [12] suggested using oleyl alcohol-water partition coefficients as a reference system for evaluating partitioning of drug in biological system. Moreover, since many NH and OH groups are present in enzymes and

membranes, it is not surprising that alcohol-water system give better correlations than aprotic solvent-water systems and thus have become more widely used as extrathermodynamic reference systems.

6.2. RELATIONSHIPS BETWEEN GAS TO SOLVENT AND GAS TO WATER PARTITIONS, WATER TO SOLVENT PARTITIONS, AND SOLUBILITIES

The gas to water partition coefficient, Kw, is given by:

$$K_W = \frac{\text{conc of solute in water}}{\text{conc of solute in the gas phase}} \tag{6.20}$$

If the concentrations in each phase are given in $mol \cdot dm^{-3}$ (or $mol \cdot L^{-1}$) then Kw is dimensionless and is equivalent to the Ostwald solubility coefficient, Lw. A similar equation leads to the gas to solvent partition coefficient, Ks (or Ls):

$$K_S = \frac{\text{conc of solute in solvent}}{\text{conc of solute in the gas phase}} \tag{6.21}$$

An alternative measure of the partition between water, or a solvent, and the gas phase is the Henry's Law constant, Hw or Hs. The former is defined as:

$$H_W = \frac{\text{amount of solute in the gas phase}}{\text{amount of solute in water}} \tag{6.22}$$

It can be seen that Hw is the inverse of Kw, and Hs will be the inverse of Ks. However, care has to be taken over units, because there are a number of ways of expressing Hw and Hs. The amount in the gas phase is usually measured as a partial pressure, VP, this can be in atm., mmHg or Pascal (1 atm = 760 mmHg = 101325 Pa). The amount in solution can be in $mol \cdot dm^{-3}$ or $mol \cdot m^{-3}$ or mol·fraction, so that there are numerous combinations of units that can be used for Hw and Hs. It is also possible to measure the amount in the gas phase (Cg) in $mol \cdot dm^{-3}$, but this is rare in studies of Hw or Hs.

One particular relationship between Ks and Hs is:

$$Ks \cdot Hs = R \cdot T \cdot \left(\frac{1000 \cdot \rho}{M} \right) \tag{6.23}$$

Here, Hs is in atm·(mol·fraction)$^{-1}$ and then ρ and M are the solvent (or water) density at temperature T and the molecular weight, respectively. R is the gas constant, whose units must be compatible with the units of Hs. If the amount in the gas phase is measured in atm then R = 0.08205 atm·K^{-1}·mol^{-1}. Then for these units, at T = 298.15K,

$$Ks \cdot Hs = 24.4632 \cdot \left(\frac{1000 \cdot \rho}{M} \right) = Q \tag{6.24}$$

Table 6.2
A few values of Q at 298.15K

Solvent	**Q**
Water	1354.0
Methanol	600.57
Cyclohexane	224.94
Hexadecane	83.17

Note that for olive oil, Ks (Ls) values are often reported at 310.15K and then Q = 25.8183 where as for olive oil at 298.15 K, Q = 24.9746

If Hw is in Pa·mol·m^{-3} as is often reported, then at 298.15K:

$$\text{Log Hw in Pa·mol·m}^{-3} - 3.394 = -\log Kw \tag{6.25}$$

The conversion between gas concentration in atm and in mol·dm $^{-1}$ at 298.15K is given by

$$\text{Log Cg} = \log \text{VP/atm} - 1.38851 \tag{6.26}$$

The usefulness of Kw and Ks is that they are related to the corresponding partition coefficient,

$$P = \frac{K_S}{K_W} = \frac{\text{conc. solute in solvent}}{\text{conc. solute in water}} \quad \text{or} \quad \log P = \log Ks - \log Kw \tag{6.27}$$

Thus if any two quantities in Eq. (6.27) are known, the third can be deduced. Eq. (6.27) applies to 'hypothetical' partitions as well as to real partitions, so that the solvent in Eq. (6.27) can be methanol or ethanol. Note also that for partition between water and solvents that take up amounts of water when saturated, such as octanol, the quantities P and Ks refer to water saturated octanol. They may not be the same for dry octanol.

Values of H or K can be obtained directly by measurement of the amount of solute above a given solution. For sparingly soluble solutes, nearly always solids, they can be obtained from the saturated vapour pressure (or concentration) and the concentration of the saturated solution, that is the solubility S. Then

$$K_W = \frac{S_W}{C_g(sat)} \text{ and } K_S = \frac{S_S}{C_g(sat)} \tag{6.28}$$

From Eq. (6.28) it is evident also that for solutes that are not too soluble

$$P = \frac{S_S}{S_W} \tag{6.29}$$

or

$$\log P = \log Ss - \log Sw \tag{6.30}$$

6.3. EXPERIMENTAL METHODS

The most extensive and useful partition coefficient data were initially obtained by simply shaking a solute with two immiscible solvents and then analysing the solute concentration in one or both phases; this was the method used in this work. This process is time consuming, tedious, prone to emulsion problems, and requires relatively large

amounts of pure compounds. In addition, there are other important practical considerations mentioned elsewhere [13].

An indirect method of obtaining partition coefficients in difficult cases, such as for very hydrophobic solutes, is by the measurement of solubilities in water, S_W, and a given solvent, S_S. The ratio S_S/S_W is then taken as the water-solvent partition coefficient [6,14].

On the other hand, recently there has been an increase in the use of separation methods such as chromatographic and electrophoretic methods. These require very little material (which does not have to be pure), are fast compared to traditional methods, and are relatively easy to automate [13]. These methods are based on the construction of a correlation model, Eq. (6.31), between a retention property characteristic of the solute, such as log k, and the separation system for a training set of solutes with known log P values. Then further measurements of the retention property in the separation system can be used to estimate log P values for other compounds:

$$\log P = p + q \log k \tag{6.31}$$

In Eq. (6.31) k is the retention factor (or retention factor resulting from the extrapolation of log k values to zero organic solvent for binary mobile phase mixture in liquid chromatography, k_W), and p and q are correlation coefficients (or fitting constant).

Some other techniques have been used to obtain partition coefficients [6] and recently further attention has been directed to the determination of the octanol-water partition coefficient [13], due to its interest in chemistry and biochemistry.

6.4. REFERENCES

1. Berthelot, Junglefleish, Ann. Chem. Phys. 4 (1872) 26
2. W. Nernst, Z. Phys. Chem. 8 (1891) 110
3. H. Meyer, Arch. Expetl. Phatol. Pharmacol. 42 (1899) 110
4. E. Overtono, *Studien uber die Narkose*, Fischer, Jena, Germany, 1901
5. J. Sangster, (1997) *Octanol-Water Partition Coefficients, Fundamentals and Physical Chemistry*, Wiley, Chichester
6. A. Leo, C. Hansh, D Elkins, Chem. Rev. 71 (1971) 525

7. A. Avdeef, (1996) Assessment of distribution-pH profiles. In V. Pliska, B. Testa, H. van de Waterbeemd (eds.), *Lipophilicity in Drug Action and Toxicology*, VCH Publishers, Weinheim, pp. 109-139
8. C. Hansch, A. Leo, D. Hoelkman, (1995) *Exploring QSAR: Fundamentals and Applications in Chemistry and Biology*, American Chemical Society, Washington, DC
9. Y. D. Lei, F. Wania, W. Y. Shiu, D. G. B. Boocock, J. Chem. Eng. Data 45 (2000) 738
10. H. Frank, M. Evans, J. Chem. Phys. 13 (1945) 507
11. M. H. Abraham, Chem Soc. Rev. 22 (1993) 73
12. K. H. Meyer, H. Hemmi, Biochem. Z., 277 (1935) 39
13. S. K. Poole, C. F. Poole, J. Chromatogr. B 797 (2003) 3
14. M. H. Abraham, M. Amin, A. M. Zissimos, Phys. Chem. Chem. Phys., 4 (2002) 5748

Chapter 7 A Brief Introduction to Statistical Analysis

7.0. INTRODUCTION

A classical multiple linear regression analysis (MLRA) is used to generate the coefficients in the Abraham's solvation equation. This is a common technique in statistics where the response or dependent variable (y), is linearly correlated to two or more predictor variables or independent variables (X), to produce equation coefficients (b_1, b_2, ..., b_k) specific to the data set under analysis [1].

$$y = b_0 + b_1 \cdot X_1 + b_2 \cdot X_2 + \ldots + b_k \cdot X_k \quad (7.1)$$

Once the coefficients are known, it is possible to predict further values of the dependent variable, y, based on new values of the predictor variables. The accuracy of the predicted variable, y, depends on the degree of scatter in the data.

Another statistical analysis carried out in this work is the Paired t-test. Using this test it is possible to detect whether or not there is a significant difference between the results obtained by two different methods caused by the introduction of systematic error in the measure by one of them.

The last statistical analysis used in this work is the Principal Component Analysis. Using this analysis it is possible to identify, in a large set of data that consists of several variables, the scatter of the data and the existence of small groups of data in the data set, that is, to evaluate the similarity between individual data.

7.1. MULTIPLE LINEAR REGRESSION ANALYSIS

7.1.1 Linear least squares

There are several methods for determining the line that best describes the correlation between the response and the predictor variables, of which 'least squares' fitting is the most commonly used. This method is based on minimizing the deviations in the y-direction between the experimental points and the calculated line, that is minimizing the

'sum of the square of the residuals', SS_E,

$$\min = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{7.2}$$

$(y_i - \hat{y}_i)$ represents the residuals of the model, where the y_i value is the i^{th} observed response and the $\hat{y}_i$ value is the i^{th} response calculated from the regression line, that is the fitted response. The residuals provide a measure of the closeness of agreement of the observed and the fitted responses; hence, they provide a measure of the adequacy of the fitted model. Small residuals are one important indicator of the adequacy of the regression fit [2].

7.1.2 Interpreting the results

7.1.2.1. The regression table

Normally, in a MLR analysis output the regression table includes the *coefficient constants*, (b_1, b_2, ..., b_k), the *standard error of the coefficient*, SE Coeff, the *T-value* for the variable, T, and the coefficient value for P, *P-value*. The following is a brief description of each statistical parameter.

Coefficients constants (b_1, b_2, ..., b_k) represents the estimated change in the mean response per unit increase in X when all other variables are held constant.

The *standard error of the estimated coefficient*, SE Coeff, is the estimated standard deviation of the coefficient. It measures the precision of the estimated coefficients. The smaller the standard error, the more precise the estimate. The formula for coefficients in multiple regression in matrix terms is

$$\text{SE Coeff} = \sqrt{((\mathbf{X}'\mathbf{X})^{-1}s^2)} \tag{7.3}$$

with **X** as the matrix of variables and s as the standard deviation of the error in the model, see below.

The *T-value* for the variable is calculated by Eq. (7.4). The bigger the absolute value of T, the more likely the variable is significant.

$$T = \frac{\text{Estimated Coefficient}}{\text{SE Coeff}} \tag{7.4}$$

The *T-value* is used to calculate the coefficient value for P, *P-value*, which is the key figure of the estimated parameters in the regression table. This helps to say whether or not the association between the response and the variables is statistically significant. If the P-value is smaller than a level of significance (α-level) selected, the association is statistically significant. A commonly used α-level is 0.05 (or 5%).

Following the regression table, some other parameters necessary to interpret the results are used in the MLRA output. Those parameters are:

The *standard deviation* of the error in the model, SD, measures the overall error associated with the regression line and is estimated from the residuals between the y value and the corresponding value calculated from the fitted line,

$$SD = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-p-1}} \tag{7.5}$$

SD is dependant upon the number of samples, n, in the data set and also upon the number of variables, p, used in the model. The standard deviation has the same units as the property being measured. Standard deviation is an estimate of the variance about the regression, i.e. about the fitted line.

The *correlation coefficient*, R^2, gives a measure of the success of the correlation of the dependent variable, y, against the independent variables, X, and is given by Eq. (7.6),

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{7.6}$$

where the $\bar{y}$ value is the mean response. R^2 is a measure of how closely the data set fits the relationship given by the MLRA and can assume values between 0 and 1. A value of 1 indicates that the data set is explained by the correlation equation perfectly, while a value of zero means there is no correlation between the data set and the model.

Neither the coefficient of determination nor the usual statistics-for example, t test, F test and P-value that accompany the regression model provide insight into the quality and the potential influence of individual observations on the estimates. The *predictive* r^2 can be used as the criterion for evaluating the power of predicting of an equation.

The *prediction sum of squares*, PRESS, allows assessing the model's internal self-consistency. PRESS is the sum of squares of the prediction error, and it is calculated by the following formula

$$\text{PRESS} = \sum_{i=1}^{n}\left(\frac{e_i}{1-h_i}\right)^2 \tag{7.7}$$

where e_i is the residual and h_i is the i^{th} diagonal element of the $\mathbf{X(X'X)^{-1}X'}$. In general, the smaller the PRESS value, the better the model's self-consistency.

The *predictive* r^2, r^2(pred), indicates how well the model predicts responses for observations within the data set, whereas r^2 indicates how well the model fits the data. r^2(pred) is given by Eq. (7.8),

$$r^2(\text{pred}) = 1 - \frac{\text{PRESS}}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{7.8}$$

The predictive r^2 can prevent overfitting the model, that is, fitting the model too closely to the data in the current data set, but it is not useful for assessment of the predictive power of a model. Predicted r^2 ranges from 0 to 1, larger values of predicted r^2 suggest models of greater self consistency.

7.1.2.2. The ANOVA table

Having obtained the values of the model parameters that give the best representation of the data, the next stage is to assess how well the chosen model actually fits the data.

The first step is to define a suitable measure of variation for which a partitioning into components due to assignable causes and to random variation can be accomplished [2]. While there are many measures that could be used, the *total sum of squares*, SS_T, is the most common,

$$SS_T = \sum_{i=1}^{n}(y_i - \bar{y})^2 \tag{7.9}$$

This represents the total sum of squares adjusted for the overall average by subtracting it from each individual response. SS_T is the total variation in the data. The partitioning of the above total sum of squares into components for the assignable causes and for random (uncontrolled experimental) variation yields the following equations (7.11) and (7.12),

$$SS_T = SS_R + SS_E \tag{7.10}$$

$$SS_R = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 \tag{7.11}$$

$$SS_E = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{7.12}$$

where SS_R is the *sum of squares regression* which indicates the portion of variation explained by the model, while the *sum of squares error*, SS_E, is the portion not explained by the model and is attributed to the error. If the fit of the model is good, SS_E is expected to be small compared with SS_R, and this forms the basis of a significance test for the regression [3].

The significance of the regression with respect to the measurement error is assessed by the *F-statistic*, which uses the ratio of the *mean square regression,* MS_R, to the *mean square error*, MS_E, see Eq. (7.13). A significant ratio means that the amount of variation in the data due to the underlying model is greater than could be explained by the variation expected as a result of the measurement noise [3]. In other words, the model and its calculated parameters are reasonable and likely to be true.

$$F = \frac{MS_R}{MS_E} = \frac{\frac{SS_R}{DF_R}}{\frac{SS_E}{DF_E}} = \frac{\frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{p}}{\frac{\sum_{i=1}^{n}(y_i - \hat{y})^2}{n-p-1}} \tag{7.13}$$

The column labelled *P* in the ANOVA table is the significance value and represents the probability of getting an F-value at least as large as the one calculated by a random chance.

The value of 0.000 found in many outputs means that the P-value when rounded is smaller than 0.001 or 0.1%. The value of < 0.001 means that there is less than 1 chance in 1000 of getting such a large value [4]. The confidence level and the significance level are related in that the confidence level = $100 \cdot (1\text{-}P)$.

7.1.2.3. Lack of fit

The ability to determine when a regression fit adequately models a response variable is greatly enhanced when repeat observations for several combinations of the predictor variables are available [2]. The error sum of square can then be partitioned into a component due to *pure error*, SSE_P, and a component due to *lack-of-fit* error, SSE_{LOF}.

$$SS_E = SSE_P + SSE_{LOF} \tag{7.14}$$

The pure error sum of squares, SSE_P, is calculated using the response of the repeat observations:

$$SSE_P = \sum_{i=1}^{m}\left(\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i\bullet})^2\right) \tag{7.15}$$

In Eq. (7.15) a double subscript has been introduced for clarity. It is only in the calculation of SSE_P that this is needed. It is assumed that there are m combinations of values of variables with repeat observations. For the i^{th} combination, n_i repeats are available, the average of which is denoted by $\bar{y}_{i\bullet}$. The total number of degrees of freedom for this estimate of error is

$$f_P = \sum_{i=1}^{n}(n_i - 1) \tag{7.16}$$

where $q = n_1 + n_2 + \cdots + n_m$. The lack-of-fit error sum of squares, SSE_{LOF}, is obtained as the difference between SS_E and SSE_P, with f_{LOF} = (n - p - 1) - (q - m) degrees of freedom.

If the lack-of-fit test is statistically significant by its F-statistics (ratio of lack of fit mean square to pure error mean square), $F = MSE_{LOF} / MSE_P$, the model being used

is wrong for the data and there is no point in trying to get any further information about the model from the data. Instead, a better model should be sought.

7.2. PAIRED T-TEST

It frequently happens that two methods of analysis are compared by studying test samples containing different amounts of analyte, as will be done in Chapter 14.

As always there is the variation between the measurements due to random measurement error. In addition, differences in the samples and differences between the methods may also contribute to the variation between the measurements. In order to overcome this difficulty the differences, d, between each pair of results by the two methods can be used instead of the y_i value of each tested sample. If there is no difference between the two methods then these differences are drawn from a population with mean $\mu_d = 0$. In order to test this, it must be analysed whether $\bar{d}$ differs significantly from 0. A 95% confidence interval for the average difference in conversion efficiency is calculated as

$$\bar{d} \pm t_{n-1,0.05} \frac{s_d}{\sqrt{n}} \tag{7.17}$$

where n is the number of paired comparisons. The standard deviation of the differences, s_d, is given by the following formula

$$s_d = \sqrt{\frac{\sum_{i=1}^{n} (d_i - \bar{d})^2}{n-1}} \tag{7.18}$$

The confidence interval is a range of likely values for μ_d. Since the true value of μ_d is not known, the confidence interval allows one to guess its value based on the sample data. The mean sample difference provides an estimate of μ_d, and the standard deviation of the sample difference s_d, is used to determine how far off the estimate might be. In general, the proportion of intervals that include μ_d is equal to 1 minus the chosen α-level.

Commonly, data are analysed with an α-level of 0.05, therefore a 95% (or 0.95) confidence interval is constructed. This interval indicates that, based on the sample data, μ_d

is greater than or equal to $\bar{d} - t_{n-1,0.05}\frac{s_d}{\sqrt{n}}$ and less than or equal to $\bar{d} + t_{n-1,0.05}\frac{s_d}{\sqrt{n}}$. If the reference value of 0 is not within the confidence interval, it can be concluded, with a 95% confidence, that μ_d is not 0. In other words, the methods are different and cannot be compared or, there is a systematic error in either one of the two methods. Any α-level greater than 0% and less than 100% can be chosen.

The paired *t-test* gives two statistics that can be used to conduct a test of the mean difference: a *T-value* and a *P-value*. The T-value is not very informative by itself, but it is used to determine the P-value. The P-value indicates just how likely it is that the same distribution of differences would be obtained, with its particular mean and standard deviation, if $\mu_d = 0$ is true.

The P-value required for the t-test must be decided before the test is conducted. The value chosen is called again the α-level. If the P-value is less than or equal to this α-level, then it can be concluded that μ is not equal to the reference value. And it can be concluded that the methods are different and cannot be compared, probably because there is a systematic error in one of the methods or just because the two methods do not perform equally.

7.3. PRINCIPAL COMPONENTS ANALYSIS

Principal Component Analysis (PCA) is a pattern recognition method. A very effective method for the analysis of quantitative data (continues or discrete) which are presented under the form of tables of n observations / k variables. This method allows the visualization and rapid analysis of the connection between the k variables, and the visualization and analysis of the n observations initially described by k variables in a graphic of two or three dimensions, built in a way that the dispersion between the data is as preserved as possible, by building a set of p factors (principal components) uncorrelated ($p \leq k$) that can be used later by other methods (for example: regression analysis). This graphic allows, as mention in the introduction, the identification of tendencies in the n observations.

PCA first finds the set of orthogonal eigenvectors of the correlation or covariance matrix of the variables. Then, the matrix of principal components (**C**) is the product of the

eigenvector matrix (**A**) with the normalized matrix of independent variables (**X**), i.e. independent variables divided by their own variance [5, 6].

$$\mathbf{C} = \mathbf{A} \cdot \mathbf{X} \tag{7.19}$$

where

$$C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1k} \\ c_{21} & c_{22} & \cdots & c_{2k} \\ \vdots & \vdots & & \vdots \\ c_{p1} & c_{p2} & \cdots & c_{pk} \end{pmatrix} ; A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{pmatrix} ; X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{pmatrix}$$

7.3.1 Eigenanalysis

The *eigenvectors*, which are comprised of coefficients corresponding to each variable, are the weights for each variable used to calculate the principal components scores. These *eigenvectors* are the *p* rows in the matrix **A**. The principal components *scores* are the linear combinations of the original variables that account for the variance in the data. The matrix **A** is, in fact, the matrix of eigenvectors associated with the matrix of eigenvalues [7]. The *eigenvalues* are the variances explained of the principal components **C**. These values are found in the *k* x *k* matrix called the eigenvalue matrix or 'loading' matrix (Λ), which has the following form

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_k \end{pmatrix}$$

The first principal component accounts for the largest percent of the total data variation. The second principal component accounts the second largest percent of the total data variation, and so on. The goal of principal components is to explain the maximum amount of variance with the fewest number of components [6].

The *proportion* of variance explained by the i^{th} principal component is given by

$$\frac{\lambda_i}{\lambda_1 + \lambda_2 + ... + \lambda_n} \tag{7.20}$$

where λ_i is the i^{th} eigenvalue. The *cumulative proportion* of variance explained by the first i principal components is obtained as follows

$$\frac{\lambda_1 + \lambda_2 + ... + \lambda_i}{\lambda_1 + \lambda_2 + ... + \lambda_n} \tag{7.21}$$

where λ_i is the i^{th} eigenvalue.

7.3.2 Analysis of the eigenvalue and score plots

The *eigenvalue plot* shows the eigenvalue associated with a principal component versus the number of the component. This plot is used to judge the relative magnitude of eigenvalues. If the model is good, the eigenvalue associated in this plot to the first principal component must be very large comparing with the rest of eigenvalues. The sharp jumps in these graphs, when they take place, indicate the approximate rank of the data matrix and, therefore, allow the establishment of the number of dimensions necessary to model, with a small or acceptable error, the original data matrix (**X**).

In these *score plots*, one normally can distinguish a number of clusters of samples [7]. They show the scores for the second principal component (y-axis) versus the scores for the first principal component (x-axis). It is in this plot, where one can observe if there is a tendency in the observations, either because they are found in different groups, or because they are arranged in some unexpected way, indicating the existence of at last one variable that really makes a different over the others.

7.4. TRAINING SET AND TEST SET

The most satisfactory way to assess the predictive power of an equation is to divide the total data sets into a representative training set, or fitting set, and a test set. These two data sets are chosen in this work by an alternative space-filling technique developed by Kennard and Stone [8]. They proposed a sequential method that should cover the whole descriptor space of the total set. The procedure consists of selecting as the next sample (candidate object) the one that is most distant from those already selected objects

(calibration objects). As starting points either the two objects that are most distant from each other are selected, or preferably, the one closest to the mean. From all the candidate points, the one is selected that is furthest from those already selected and added to the set of calibration points. To do this, the distance from each candidate point i_0 to each point i which has already been selected is measured and which is the smallest is determined through

$$(\min_{i}(d_{i,i_0})) \tag{7.22}$$

From these the one for which the distance is maximal is selected,

$$d_{selected} = \max_{i_0}(\min_{i}(d_{i,i_0})) \tag{7.23}$$

These two sets of data are used in the following way. The training set is used to set up either Eq. (7.24) or Eq (7.25) (already described in Chapter 4), depending on the solute parameter, SP, under study, and thus to obtain the coefficient system of the equation. Once this is done, the equation obtained is used to predict the SP values for the compounds in the test set that have not been used to construct the equation, by introducing their solute descriptors. As a result, the differences between the calculated values and the experimentally observed values can be compared, and a conclusion on the predictive power of the equation can be achieved.

$$SP = c + e \cdot \mathbf{E} + s \cdot \mathbf{S} + a \cdot \mathbf{A} + b \cdot \mathbf{B} + v \cdot \mathbf{V} \tag{7.24}$$

$$SP = c + e \cdot \mathbf{E} + s \cdot \mathbf{S} + a \cdot \mathbf{A} + b \cdot \mathbf{B} + l \cdot \mathbf{L} \tag{7.25}$$

7.5. REFERENCES

1. L. Eriksson, E. Johansson, Chemom. Intell. Lab. Syst. 34 (1996) 1
2. R. L. Manson, R. F. Gunst, J. L. Hess, (2003) *Statistical Design and Analysis of Experiments*, 2nd ed., Wiley-Interscience, New Jersey
3. M. Porter, (2000) Analysis of Variance. Understanding and modelling variability. In R. L. Tranter (ed.), *Design and Analysis in Chemical Research*, Sheffield Academic Press Ltd., Sheffield, England, p. 279

4. N. R. Draper, H. Smith, (1998) *Applied Regression Analysis*, 3rd ed., Wiley-Interscience, New York
5. R. Kellner, J.-M. Mermet, M. Otto, M. Valcarcel, H. M. Widmer (Eds.), (2004) *Analytical Chemistry: A modern approach to Analytical Science*, 2nd ed., Wiley-VCH
6. I. T. Jolliffe, (2002) *Principal Component Analysis*, 2nd ed., Springer-Verlag New York, Inc.
7. D. L. Massart, B. G. M. Vandeginste, L. M. C. Buydens, S. De Jong, P. J. Lewi, J. Smeyers-Verbeke (Eds.), (1997) *Handbook of chemometrics and qualimetrics, Part A*, Elsevier Science, Amsterdam
8. R.W. Kennard, L. A. Stone, Technometrics 11 (1969) 137

Chapter 8 Aims of the Present Work

The presence of volatile organic compounds (VOCs) in the environment can lead to health effects, such us sensory irritation (nasal pungency and eye irritation). Describing the physicochemical characteristics of a volatile organic compound (VOC), in terms of dipolarity, hydrogen bonding capacity and lipophilicity, can help to interpret the mechanism of the interaction of VOCs with chemosensory receptors areas, which is the final goal of this project.

In previous studies by Abraham and co-workers, it has been shown that nasal pungency thresholds (NPTs), odour detection thresholds (ODTs) and eye irritation thresholds (EITs) for VOCs can be correlated with VOC physicochemical properties through the general Eq. (8.1), that refers to the transport of compounds from the gas phase to a condensed phase. The specific equations for ODT, NPT and EIT are given as Eq. (8.2 – 8.4).

$$SP = c + e \cdot \mathbf{E} + s \cdot \mathbf{S} + a \cdot \mathbf{A} + b \cdot \mathbf{B} + l \cdot \mathbf{L} \quad (8.1)$$

Thus, the first aim of this project is to obtain updated quantitative structure activity relationships (QSARs) for log (1/NPT) and log (1/ODT) on the lines of Eq (8.1 – 8.4).

$$\log (1/ODT) = - 5.27 + 0.51\,\mathbf{E} + 1.96\,\mathbf{S} + 1.48\,\mathbf{A} + 1.53\,\mathbf{B} + 0.723\,\mathbf{L} \quad (8.2)$$

$$\log (1/NPT) = - 7.89 + 0.20\,\mathbf{E} + 1.32\,\mathbf{S} + 2.71\,\mathbf{A} + 1.52\,\mathbf{B} + 0.823\,\mathbf{L} \quad (8.3)$$

$$\log (1/EIT) = -7.89 - 0.38\,\mathbf{E} + 1.87\,\mathbf{S} + 3.78\,\mathbf{A} + 1.17\,\mathbf{B} + 0.785\,\mathbf{L} \quad (8.4)$$

QSARs developed in this way can only be used to predict NPT and ODT values for VOCs if solute descriptors are available. These descriptors are obtained through the use of various thermodynamic measurements, i.e., water to solvent partition coefficients and gas chromatographic retention data. The second aim of the project is to obtain descriptors for environmentally important VOCs, using both data determined in this work, and literature data where available.

The third aim of the project is to study the possibility that NPT values could be used to estimate 'selective' effects in ODTs, thus providing a method to assess the weight of the remaining 'specific' (VOC-receptor) effects.

The fourth aim is to study concentration-detection, i.e., detectability, functions for the human olfactory detection of several chemical vapours of diverse chemical structures measured. The functions cover the complete peri-threshold range from chance to almost perfect detection and provide much more information than what a single "threshold value" can convey. These odour and nasal pungency functions have been modeled with a uniform approach: a simplified sigmoid equation, $y = 1 / (1 + e^{[-(x-C)/D]})$, and have been correlated with the set of up to five physicochemical descriptors taken from the Abraham general equation described above.

The fifth and final aim of this work is to study the existence of "cut-off", i.e., a point where a homolog fails to stimulate even at vapor saturation, suggested by previous studies on ocular and nasal chemesthetic thresholds along and across homologous chemical series. In attempt to explain this "cut-off" the possibility of there being some critical molecular dimension aspect, beyond which ocular chemesthesis is not evoked, has been explored.

Chapter 9 Determination of the Water/Octanol Partition Coefficient as the log P value for Fluorescein and Diacetylfluorescein

9.0. AIM

The main purpose of this work was to find the water/octanol partition coefficient as the log P value (log P_{oct}) for fluorescein and diacetylfluorescein. This work was carried out in order to obtain expertise in the practical determination of partition coefficients.

The partition coefficient P is defined as the ratio of the final concentrations (assuming that only the neutral species is considered):

$$P = \frac{Cs^f}{Cw^f} \tag{9.1}$$

where Cs^f is the final concentration of the solute in the solvent and Cw^f is the final concentration of the solute in the aqueous solution. When the solvent is octanol, the partition coefficient is denoted as P_{oct}.

Almost all of the log P_{oct} values found in the literature for fluorescein and dicetylfluorescein were calculated values. The only found experimental value by van Bree et al. [1] was obtained for fluorescein at pH 6.8 and does not refer to the neutral form, see Figure 9.1 and Table 9.8, later.

I FLUORESCEIN

9.1. INTRODUCTION

Fluorescein is an organic compound whose structure depends on the pH. Five forms can exist in aqueous solution: cationic, neutral, lactone, anionic and dianionic forms, making its absorption properties strongly pH dependent. In practice, the lactone form predominates over the neutral form, and so only four forms need to be considered. The chemical structures of the fluorescein protolytic forms are shown in Figure 9.1.

As mentioned earlier, the aim of this work is to determine the water/octanol log P value for the lactone species of fluorescein. This species reaches its maximum

molar ratio in acidic media, at pH around three as can be seen from Figure 9.2 [3]. Therefore, in order to study the lactone species in aqueous solution, experiments must be carried out at near pH three. The SMILES notation for the lactone form of fluorescein is Oc5ccc2c(Oc1cc(O)ccc1C23OC(=O)c4ccccc34)c5.

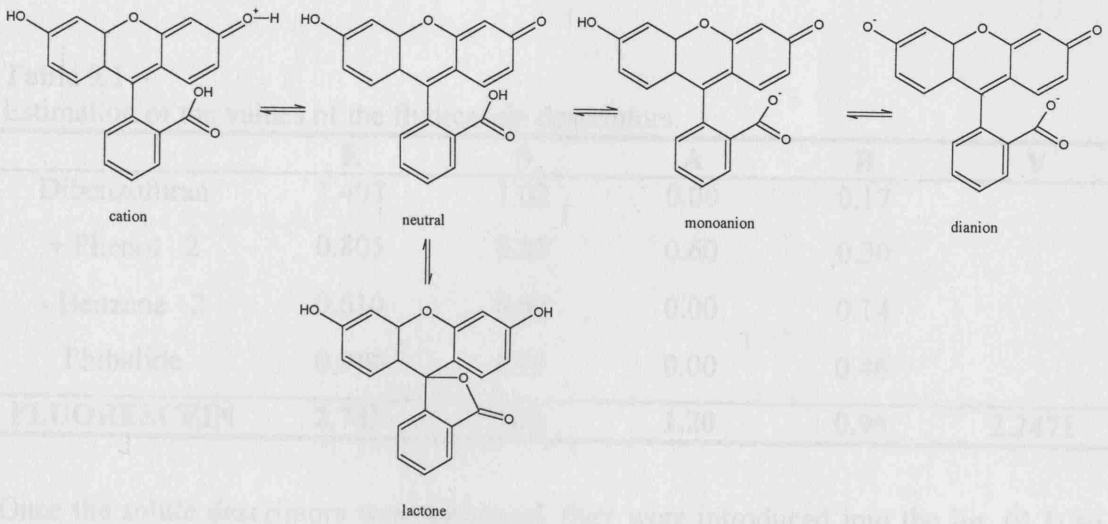


Figure 9.1. Ionization equilibria of fluorescein.

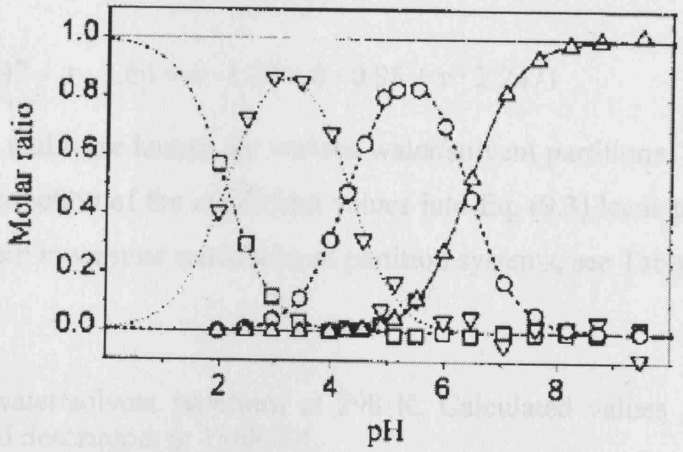


Figure 9.2. Calculated concentration of the cation (□), lactone species (▽), monoanion (○), and dianion (△). Source: Ref. 4

9.2. METHOD

9.2.1 General analysis

After several years work, Abraham *et al.,* developed the Abraham general solvation Equation (Eq. 9.2), which was dealt with in detail in Chapter 4.

$$SP = c + e \cdot \mathbf{E} + s \cdot \mathbf{S} + a \cdot \mathbf{A} + b \cdot \mathbf{B} + v \cdot \mathbf{V} \quad (9.2)$$

As a first step, the descriptors for fluorescein were estimated from known descriptors for dibenzofuran, phenol, benzene, and phthalide by simple addition. Results are shown in Table 9.1.

Table 9.1
Estimation of the values of the fluorescein descriptors.

	E	S	A	B	V
Dibenzofuran	1.407	1.02	0.00	0.17	
+ Phenol · 2	0.805	0.89	0.60	0.30	
- Benzene · 2	0.610	0.52	0.00	0.14	
Phthalide	0.950	1.90	0.00	0.46	
FLUORESCEIN	**2.747**	**3.66**	**1.20**	**0.95**	**2.2471**

Once the solute descriptors were estimated, they were introduced into the Eq. (9.2) so that values of log P could be predicted for fluorescein in various water/solvent partition systems. As a result, Eq. (9.3) was obtained,

$$\log P = c + e \cdot 2.747 + s \cdot 3.66 + a \cdot 1.20 + b \cdot 0.95 + v \cdot 2.2471 \quad (9.3)$$

where *c*, *e*, *s*, *a*, *b* and *v* are known for various water/solvent partitions. These are given in Table 9.2; introduction of the coefficient values into Eq. (9.3) leads to predictions of log P for fluorescein in various water/solvent partition systems, see Table 9.2.

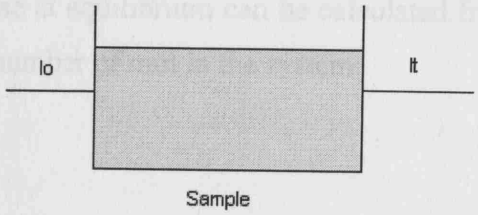
Table 9.2
Coefficients for water/solvent partitions at 298 K. Calculated values for fluorescein, using the estimated descriptors in Table 9.1.

Solvent	c	e	s	a	b	v	Log P_{calc}
Octanol	0.088	0.562	-1.054	0.034	-3.460	3.814	3.098
Dichloromethane	0.314	0.001	0.022	-3.238	-4.137	4.259	2.152
Trichloromethane	0.327	0.157	-0.391	-3.191	-3.437	4.191	1.650
1,2-Dichloroethane	0.227	0.278	-0.167	-2.816	-4.324	4.205	2.342
Toluene	0.143	0.527	-0.720	-3.010	-4.824	4.545	0.974
Chlorobenzene	0.040	0.246	-0.462	-3.038	-4.769	4.640	1.275
Bromobenzene	-0.130	0.394	-0.280	-3.331	-4.640	4.583	1.821

9.2.2 *Determination of the Partition Coefficient*

Some molecules, such as fluorescein, are coloured and thus absorb radiation in the visible region of the spectrum. The electrons in the bonds within the molecule become excited so that they occupy a higher quantum state and in the process absorb some of the energy passing through the solution.

Absorption of light by a solute



The measurement of light absorption by a solution of molecules is governed by the Beer-Lambert Law [4], which is written as follows:

$$\log \frac{I_0}{I_t} = A = \varepsilon \cdot b \cdot c \tag{9.4}$$

where I_0 is the intensity of incident ration, I_t is the intensity of transmitted radiation; A is known as the absorbance and is a measure of the amount of light absorbed by the sample; ε is a constant known as the molar extinction coefficient and is the absorbance of a 1M solution of the analyte, b is the path length of the cell in cm, and c is the concentration of the analyte in $mol \cdot l^{-1}$.

The partition coefficient of a molecule between water and an organic solvent may be determined by shaking the organic solvent and the water layer together and determining the amount of molecule in either the aqueous or organic layer by UV visible spectrophotometry [4]. A method to calculate the partition coefficient of a solute between water and an organic solvent is as follows. If the initial concentration of the solute in water is Cw^i in units of $mol \cdot dm^{-3}$ and the volume of the water phase is Vw in units of dm^3, the number of solute mol in the initial water phase is

$$Mw^i = Cw^i \cdot Vw \tag{9.5}$$

and this is also the total number of solute mol in the system, M^t.

A known amount of solvent, Vs, is added into the aqueous solution. After equilibration, the final concentration of the solute in the solvent is Cs^f, and the number of mol of solute in the solvent will be

$$Ms^f = Cs^f \cdot Vs \tag{9.6}$$

The number of mol solute in the water phase at equilibrium can be calculated from the number of mol in the solvent, and the total number of mol in the system:

$$Mw^f = M^t - Ms^f \tag{9.7}$$

The final concentration in the water phase is then:

$$Cw^f = \frac{Mw^f}{Vw} \tag{9.8}$$

The partition coefficient is defined as the ratios of the final concentrations, assuming that only the neutral species is considered:

$$P = \frac{Cs^f}{Cw^f} \tag{9.1}$$

Now

$$Cs = \frac{Ms^f}{Vs} = \frac{M^t - Mw^f}{Vs} = \frac{M^t - Cw^f \cdot Vw}{Vs} = \frac{Cw^i \cdot Vw - Cw^f \cdot Vw}{Vs} = \frac{\left(Cw^i - Cw^f\right) \cdot Vw}{Vs}$$

Thus, Eq. (9.1) can be rewritten as:

$$P = \frac{\left(Cw^i - Cw^f\right) \cdot Vw}{Cw^f \cdot Vs} = \left(\frac{Cw^i}{Cw^f} - 1\right) \cdot \frac{Vw}{Vs} \tag{9.9}$$

The absorbance was defined as

$$A = \varepsilon \cdot b \cdot c \tag{9.4}$$

So if the absorption of the water phase at the initial concentration and the final concentration is measured at the same wavelength and the same length cell, then $Cw^i/Cw^f = Aw^i/Aw^f$. In this case, Eq. (9.9) takes the form

$$P = \left(\frac{Aw^i}{Aw^f} - 1\right) \cdot \frac{Vw}{Vs} \tag{9.10}$$

Thus, all that is needed to obtain a value of P is the absorption measurement on the initial and final water phase, and the volumes of the water and solvent phases.

9.2.3 Experimental

The solutions with fluorescein were prepared as follows: 3000 ml water plus a little amount of fluorescein + 100 ml methanol + 3 full pasteur pipette of HCl 0.5 M. The solutions were left stirring for around two weeks with the purpose of obtaining a high concentration of fluorescein in the solutions. After two weeks the solutions were filtered to remove the excess of fluorescein.

The systems studied are described in Table 9.3. In this table the superscript **a** means the volume ratios used in the first test were estimated with the value of log P calculated, and the superscript **b** means the ratios were estimated with the value of log P observed in that test (see Table 9.4).

Table 9.3
Systems studied

Toluene[a]	**Toluene[b]**
1. 250 ml solution + 20 ml Toluene	1. 150 ml solution +100 ml Toluene
2. 250 ml solution + 20 ml Toluene	2. 300 ml solution +100 ml Toluene
3. 250 ml solution + 20 ml Toluene	3. 600 ml solution + 100 ml Toluene
Trichloromethane[a]	**Trichloromethane[b]**
1. 700 ml solution + 10 ml Trichloromethane	1. 150 ml solution + 30 ml Trichloromethane
2. 300 ml solution + 10 ml Trichloromethane	2. 200 ml solution + 20 ml Trichloromethane
3. 1000 ml solution + 10 ml Trichloromethane	3. 400 ml solution + 20 ml Trichloromethane

Chlorobenzene[a]

1. 300 ml solution + 10 ml Chlorobenzene
2. 700 ml solution + 10 ml Chlorobenzene
3. 1000 ml solution + 10 ml Chlorobenzene

Chlorobenzene[b]

1. 200 ml solution + 50 ml Chlorobenzene
2. 300 ml solution + 50 ml Chlorobenzene
3. 400 ml solution + 50 ml Chlorobenzene

1,2-Dichloroethane[a]

1. 500 ml solution + 100 ml 1,2-Dichloroethane
2. 500 ml solution + 10 ml 1,2-Dichloroethane
3. 800 ml solution + 10 ml 1,2-Dichloroethane

1,2-Dichloroethane[b]

1. 580 ml solution + 20 ml 1,2-Dichloroethane
2. 580 ml solution + 16 ml 1,2-Dichloroethane
3. 580 ml solution + 12 ml 1,2-Dichloroethane

Bromobenzene[a]

1. 200 ml solution + 25 ml Bromobenzene
2. 200 ml solution + 20 ml Bromobenzene
3. 200 ml solution + 15 ml Bromobenzene

Dichloromethane[a]

1. 400 ml solution + 40 ml Dichloromethane
2. 400 ml solution + 50 ml Dichloromethane
3. 400 ml solution + 70 ml Dichloromethane

Octanol[1]

1. 2000 ml solution + 1 ml Octanol[a]
2. 2000 ml solution + 0.9 ml Octanol[b]

[1] The solution must be saturated with octanol before adding the octanol.

The partition coefficients were determined by measurements of absorbance at *436 nm.* This wavelength was observed as the maximum wavelength (λ_{max}) for fuorescein after obtaining its spectrum at *pH 3.28* (see figure 9.3) [2]. Absorption spectra were measured on a Kontron spectrophotometer. The optical path was 5 cm and spectra were collected at a resolution of one data point per nanometer.

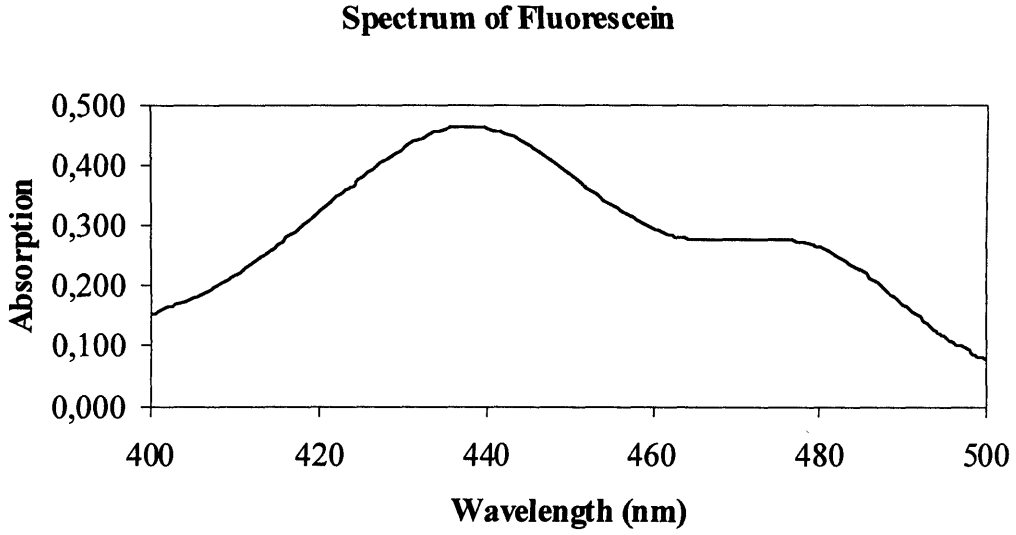


Figure 9.3. Spectrum of fluorescein at pH 3.28 in a cuvette with a path length of 5 cm.

9.3. RESULTS

Table 9.4
Results found in this work. Observed absorbance and partition coefficient.

Toluene[a]			
Test	1	2	3
Date			
06/06/03	0.591	0.524	0.531
09/06/03	0.443	0.461	0.459
10/06/03	0.446	0.454	0.444
11/06/03	0.433	0.442	0.434
12/06/03	0.451	0.461	0.462
13/06/03	0.451	0.460	0.455
P	4.109	1.877	2.224

Toluene[b]			
Test	1	2	3
Date			
16/06/03	0.406	0.423	0.411
17/06/03	0.154	0.224	0.285
18/06/03	0.152	0.226	0.288
19/06/03	0.153	0.224	0.278
P	2.480	2.615	2.562

Trichloromethane[a]			
Test	1	2	3
Date			
10/06/03	0.517	0.525	0.517
11/06/03	0.457	0.400	0.466
12/06/03	0.455	0.403	0.463
13/06/03	0.456	0.409	0.461
P	9.364	8.985	11.663

Trichloromethane[b]			
Test	1	2	3
Date			
16/06/03	0.410	0.408	0.408
17/06/03	0.136	0.210	0.283
18/06/03	0.139	0.218	0.287
19/06/03	0.138	0.225	0.285
P	9.855	8.716	8.632

Chlorobenzene[a]			
Test	1	2	3
Date			
25/06/03	0.529	0.501	0.533
26/06/03	0.461	0.468	0.498
27/06/03	0.460	0.463	0.492
30/06/03	0.469	0.466	0.497
P	4.276	5.258	7.460

Chlorobenzene[b]			
Test	1	2	3
Date			
30/06/03	0.542	0.548	0.552
01/07/03	0.289	0.340	0.387
02/07/03	0.299	0.333	0.377
03/07/03	0.287	0.329	0.372
04/07/03	0.292	0.338	0.383
P	3.425	3.728	3.561

1,2-Dichloroethane[a]			
Test	1	2	3
Date			
30/06/03	0.404	0.547	0.531
01/07/03	0.041	0.336	0.339
02/07/03	0.042	0.332	0.405
03/07/03	0.040	0.336	0.411
04/07/03	0.040	0.339	0.412
P	44.268	31.399	28.367

1,2-Dichloroethane[b]			
Test	1	2	3
Date			
14/07/03	0.747	0.727	0.718
15/07/03	0.326	0.373	0.435
16/07/03	0.319	0.369	0.437
17/07/03	0.334	0.383	0.451
18/07/03	0.341	0.392	0.459
P	38.068	34.784	31.216

Table 9.4 *(continued)*

Bromobenzene			
Test	1	2	3
Date			
18/07/03	0.672	0.672	0.672
21/07/03	0.406	0.455	0.496
21/07/03	0.397	0.457	0.482
22/07/03	0.440	0.503	0.525
22/07/03	0.428	0.483	0.509
P	4.861	4.147	4.480

Dichloromethane			
Test	1	2	3
Date			
22/07/03	0.671	0.608	0.401
23/07/03	0.174	0.122	0.044
23/07/03	0.186	0.134	0.050
24/07/03	0.216	0.158	0.063
P	24.948	27.246	38.352

Octanol			
Test 1		Test 2	
Date		Date	
28/07/03	0.432	28/08/03	0.614
20/08/03	0.254	16/09/03	0.389
20/08/03	0.239	17/09/03	0.350
21/08/03	0.229	18/09/03	0.329
21/08/03	0.231	19/09/03	0.311
22/08/03	0.205	22/09/03	0.302
26/08/03	0.194	23/09/03	0.299
27/08/03	0.194	24/09/03	0.303
28/08/03	0.204	25/09/03	0.301
P	2341.709	P	2310.816

Note that the first measurement in each test in Table 9.4 is the absorbance before adding the solvent to calculate the partition coefficient.

Table 9.5
Summary of the observed values in this work.

Solvent	log P_{obs}	n	S.D. (for n = 6)
Octanol	3.367	2	
Dichloromethane	1.472	3	
Trichloromethane	0.977	6	0.049
1,2-Dichloroethane	1.535	6	0.070
Toluene	0.409	6	0.113
Chlorobenzene	0.647	6	0.130
Bromobenzene	0.652	3	

9.4. CALCULATION OF DESCRIPTORS FOR FLUORESCEIN

To calculate the descriptors for fluorescein several procedures were used. First of all, **V** could be obtained by exact calculation and **E** could be obtained by addition as was done in Table 9.1. On the other hand, **S**, **A** and **B** needed to be calculated from the experimental log P values. To do this, the mathematical capabilities of a spreadsheet program such as Excel to optimize a problem by adjusting the variables was used. In this case, it was used to solve the set of equations for water/solvent systems given in Table 9.2 by trial and error. Note that the calculated log P values given in Table 9.2 are the estimated ones, obtained using the estimated descriptors in Table 9.1. Results of the calculation of descriptors using 'Solver' on all seven systems are shown in Table 9.6 and Table 9.7.

Table 9.6
Calculated values of log P for fluorescein by the Solver method; the corresponding obtained descriptors are in Table 9.7.The observed values are from Table 9.5.

Solvent	*c*	*e*	*s*	*a*	*b*	*v*	log P_{calc}	log P_{obs}
Octanol	0.088	0.562	-1.054	0.034	-3.460	3.814	3.333	3.367
Dichloromethane	0.314	0.001	0.022	-3.238	-4.137	4.259	1.224	1.472
Trichloromethane	0.327	0.157	-0.391	-3.191	-3.437	4.191	0.925	0.977
1,2-Dichloroethane	0.227	0.278	-0.167	-2.816	-4.324	4.205	1.569	1.535
Toluene	0.143	0.527	-0.720	-3.010	-4.824	4.545	0.344	0.409
Chlorobenzene	0.040	0.246	-0.462	-3.038	-4.769	4.640	0.544	0.647
Bromobenzene	-0.130	0.394	-0.280	-3.331	-4.640	4.583	0.962	0.652

The estimated descriptors in Table 9.1 (**S** = 3.66, **A** = 1.20 and **B** = 0.95) compare very well with the final experimental descriptors (Table 9.7), but most descriptors calculated by the PHA software are considerably different (**S** = 2.44, **A** = 1.00 and **B** = 1.27), although the value for the **E** descriptor (2.72) is almost exactly the same as the estimated value in Table 9.1 (2.75).

It is possible to extend the number of equations used in the Solver analysis through the identity, Eq. (9.11),

$$\log K_{sol} - \log K_w = \log P_{sol} \qquad (9.11)$$

Here K_{sol} is the gas to solvent partition coefficient, log K_w is the gas to water partition coefficient, and P_{sol} is the corresponding water to solvent partition coefficient. A value of log K_{sol} is assumed, and from the known log P_{sol} values in Table 9.5, it is possible to deduce the log K_{sol} values. Equations for log K_{sol} are known for all the solvents except bromobenzene, and hence six new equations can be incorporated into the Solver analysis (these equations have the **L**-descriptor instead of **V**). In addition there are available two equations for log K_w and so eight new equations can be used, at the expense of having two new variables, the **L** descriptor and log K_w. When the analysis was carried out on the 15 equations, log K_w, was assigned the 'best' value of 17.12, and the 15 equations were fitted with SD = 0.153 log units. The obtained descriptors are in Table 9.7, and are identical to those from the 7 equation set, except that **S** is a little lower.

Table 9.7
Results found for the solute descriptors using the seven equations in Table 9.5 and the total set of 15 equations.

number of equations	E	S	A	B	L	V
7	2.75	3.28	1.41	1.00		2.247
15	2.75	3.20	1.42	1.00	13.94	

9.5. CONCLUSION

In this work values of log P have been determined for the neutral species of fluorescein by the measurement of the absorbance, as explained above. In particular, the water/octanol partition coefficient, as log P_{oct} for the neutral species of fluorescein has been obtained. The log P_{oct} values found in this study can be compared to other values in the literature and these comparisons are shown in Table 9.8.

Table 9.8
Comparison between the value found in this study and the values found in the literature.

Published log P value	References
2.84 (calculated)	Clog P v2.0
3.98 (calculated)	Clog P v3.0
2.70 (calculated)	Clog P v4.0
3.81 (calculated)	Clog P v4.72

Table 9.8 *(continued)*

Published log P value	References
2.66 (calculated)	CLogP v4.92
3.01 (calculated)	PHA v3.0
2.98 (calculated)	ACD v6.0
4.51(calculated)	SPARC [a]
-0.21 (observed, pH 6.8)	Van Bree et al. [1]
3.52 (observed)	Eric Clarke[b]
3.37 (observed, pH 3.3)	This work

[a] See text. [b] The experimental Eric Clarke value was obtained by running an HPLC method using an octanol column and a little amount of the same sample of fluorescein used in this work.

The previous observed value of log Poct (-0.21) was obtained at pH 6.8; at this pH the species present are the mono-anion and the di-anion, see Figure 9.2, and so the value of log Poct is expected to be very much smaller than for the neutral species. van Bree *et al.* [1] also determined log P for the water/ethyl acetate system at pH 6.8 and found a value of -0.49, again very small and again due to the mono-anion and the di-anion.

The calculated values for log Poct obtained from various software programes are in Table 9.8; many of the calculated values differ considerably from our experimental value of 3.37 log units. This is probably a consequence of the value used in the software to estimate the contribution of the fragment containing the lactone group (see Figures 9.1 and 9.4). First, the lactone group is not too common, and the number of molecules in any training set with the lactone group will be quite small. Second, the lactone fragment is in a perpendicular plane to the rest of the molecule, see Figure 9.4, and this may affect its contribution to log Poct. Some versions of the ClogP program calculate log Poct values higher than 3.37, no doubt due to whatever value is estimated for the lactone fragment. The latest version v4.92 gives 2.66, a very low value indeed. The SPARC program [5] does not calculate log Poct for the water/octanol system but can calculate a log P value for the hypothetical partition between water and dry octanol as 4.26; from the descriptors for fluorescein and coefficients for partition into wet octanol and dry octanol it can be estimated that the SPARC calculation for wet octanol will give a value of 4.51 for log Poct.

Eric Clarke, of Syngenta, has very kindly determined log P_{oct} for fluorescein using a standard method based on HPLC capacity factors using a stationary phase

saturated with octanol. The value obtained by Eric Clarke, 3.52, is quite close to the experimental value obtained in this work.

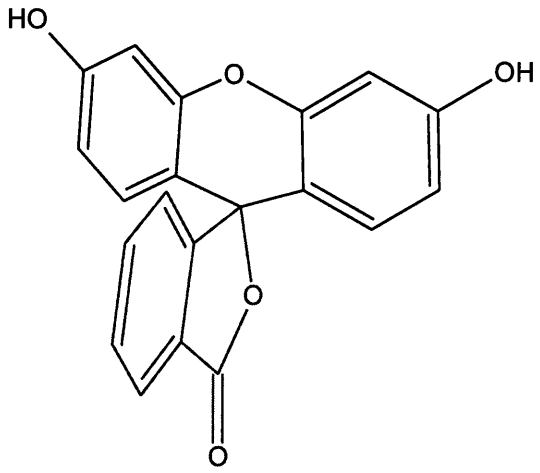


Figure 9.4. 3D Structure Optimization.

II DIACETYLFLUORESCEIN

9.6. METHOD

9.6.1 General analysis

The diacetyl derivative of fluorescein has the two phenolic groups replaced by the acetyl group, -$OCOCH_3$. The SMILES nomenclature for diacetylfluorescein is O(c5ccc2c(Oc1cc(OC(=O)C)ccc1C23OC(=O)c4ccccc34)c5)C(=O)C.

First of all, the descriptors for fluorescein, 2·phenol and 2·phenylacetate were used to estimate the values of the diacetylfluorescein descriptors by simple addition, as was done in the case of fluorescein.

Table 9.9
Estimation of the values of the diacetylfluorescein descriptors.

	E	S	A	B	V
Fluorescein	2.747	3.20	1.42	1.00	
- Phenol · 2	0.805	0.89	0.60	0.30	
+ Phenylacetate · 2	0.661	1.13	0.00	0.54	
DIACETYLFLUORESCEIN	**2.459**	**3.68**	**(0.22)**	**1.48**	**2.8421**

Once the solute descriptors were estimated, they were introduced into the Abraham Equation to predict the values of log P, which would be expected for diacetylfluorescein in various water/solvent partition systems. Note that the **A** descriptor has an estimated value of 0.22. However, there is not hydrogen bond acid present in the structure of diacetylfluorescein. Thus, the value of 0.22 for **A** is not very realistic there being no reason why it should not be zero and it is must be corrected to zero. As a result, the following equation is obtained.

$$\log P = c + e\,2.459 + s\,3.68 + b\,1.48 + v\,2.8421 \qquad (9.12)$$

where c, e, s, b and v are known for each solvent used in the present work. The values of these coefficients for some systems are given in Table 9.10, where the value of log P has been calculated for those systems using the estimated descriptors in Eq. (9.12).

9.6.2 Experimental

The solutions with diacetylfluorescein were prepared in the following way:

- Solution A: 500 ml Water + 250ml Methanol + a little amount of diacetylfluorescein.
- Solution B: 550 ml Water + 200ml Methanol + a little amount of diacetylfluorescein.
- Solution C: 600 ml Water + 150ml Methanol + a little amount of diacetylfluorescein.

The solutions were left stirring around two weeks with the purpose of getting a high concentration of diacetylfluorescein in the solutions. After these two weeks the solutions were filtered to remove the excess of diacetylfluorescein.

Table 9.10
Coefficients for water/solvent partitions at 298 K. Calculated log P (log P_{calc}) values for diacetylfluorescein in different systems, using the estimated descriptors in Table 9.9.

Solvent	*c*	*e*	*s*	*a*	*b*	*v*	log P_{calc}
Octanol	0.088	0.562	-1.054	0.034	-3.460	3.814	3.310
Tetrachloromethane	0.260	0.573	-1.254	-3.558	-4.588	4.589	3.306
Octane	0.223	0.642	-1.647	-3.480	-5.067	4.526	1.105
Hexadecane	0.087	0.667	-1.617	-3.587	-4.869	4.433	1.170
Cyclohexane	0.159	0.784	-1.678	-3.740	-4.929	4.577	1.625
Diisopropylether	0.197	0.695	-1.220	-0.238	-4.921	4.388	2.319
Dibutylether	0.184	0.817	-1.495	-0.830	-5.090	4.694	2.291

The first system studied was the water/cyclohexane system. It was studied three times. In each test, only one of the solutions (A, B or C) prepared as described above was used. Three different solutions with three different volume ratios were used in each study. The volume of the solutions studied are as shown in Table 9.11.

The partition coefficients were determined by measurements of absorbance at *453 nm.* This wavelength was observed as the maximum wavelength (λ_{max}) for diacetylfluorescein after obtaining its spectrum (see Figure 9.5). Absorption spectra

were measured on the same Kontron spectrophotometer. The optical path was 5 cm and spectra were collected at a resolution of one data point per nanometer.

Table 9.11
Solutions prepared to study the water/cyclohexane system.

Cyclohexane	Cyclohexane
1. 120 ml solution A + 30 ml Cyclohexane	1. 120 ml solution B + 30 ml Cyclohexane
2. 140 ml solution A + 30 ml Cyclohexane	2. 140 ml solution B + 30 ml Cyclohexane
3. 160 ml solution A + 30 ml Cyclohexane	3. 160 ml solution B + 30 ml Cyclohexane
Cyclohexane	
1. 120 ml solution C + 30 ml Cyclohexane	
2. 140 ml solution C + 30 ml Cyclohexane	
3. 160 ml solution C + 30 ml Cyclohexane	

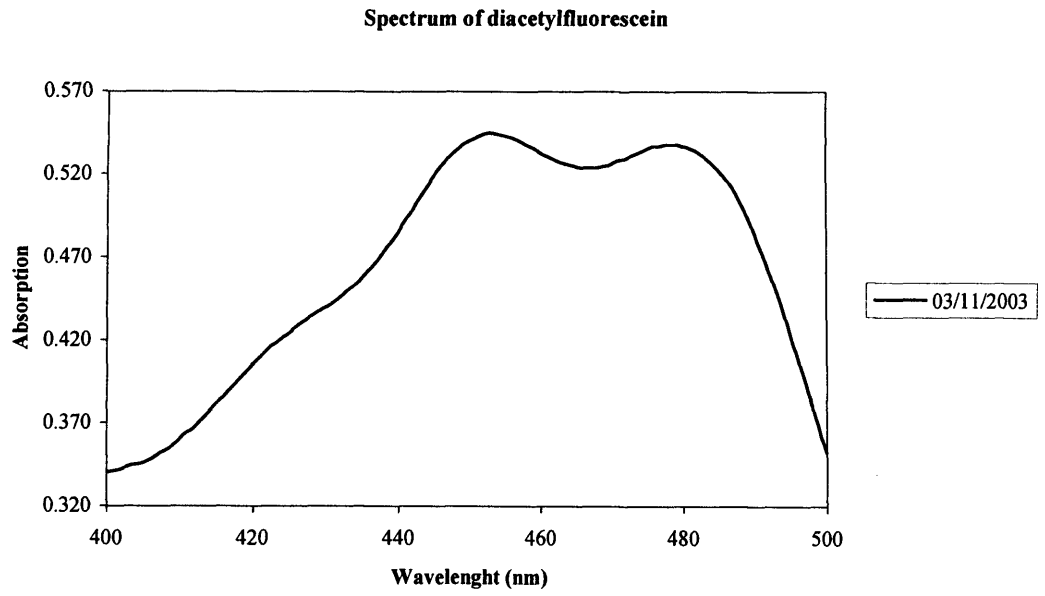


Figure 9.5. Spectrum of diacetylfluorescein in a cuvette exhibiting a path length of 5 cm.

9.7. RESULTS

The observed absorbance found for the three different solutions prepared from the A original solution are shown in Table 9.12. After observing that the absorbance increased instead of decreasing as expected, the spectra for diacetylfluorescein was taken again for both the A original solution and the prepared solution with 160 ml solution A + 30 ml cyclohexane. The results are shown in Figure 9.6.

Table 9.12
Observed absorbance obtained for the water/cyclohexane system using the solution A.

Date \ Test	1	2	3
06/11/2003	0.556	0.556	0.556
07/11/2003	0.480	0.463	0.467
08/11/2003	–	0.534	0.537

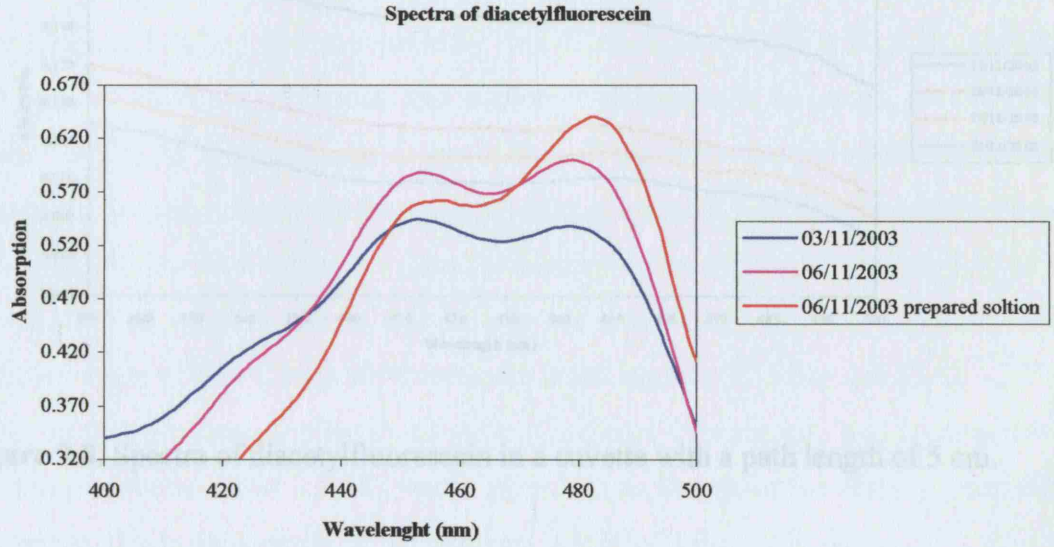


Figure 9.6. Spectra of diacetylfluorescein in a 5 cm cuvette.

As can be seen the spectrum of diacetylfluorescein changed from the first day on which it was taken, so that some possible reaction may have taken place. One cause of this could be a transesterification reaction [5] as follows

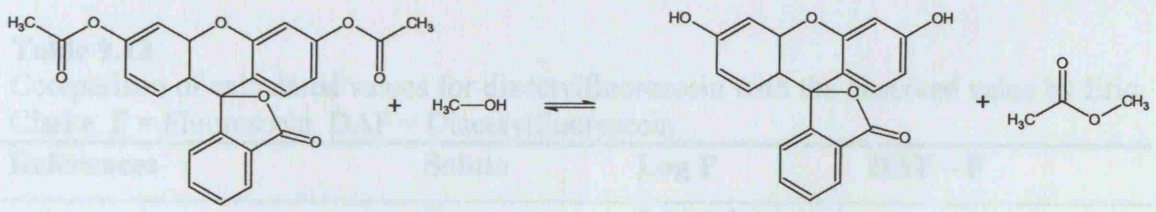


Figure 9.7. Transesterification reaction.

The experiment was then set up with acetone instead of methanol as a co-solvent with the purpose of avoiding any trans-esterification. The new solution was prepared as follows: 1000 ml Water + 70ml acetone + a little amount of diacetylfluorescein.

In this case, the solution was left stirring just one week with the purpose of getting a high concentration of diacetylfluorescein in the solutions. After one week the solution was filtered to remove the excess of diacetylfluorescein, and the spectrum was taken on four succesive days.

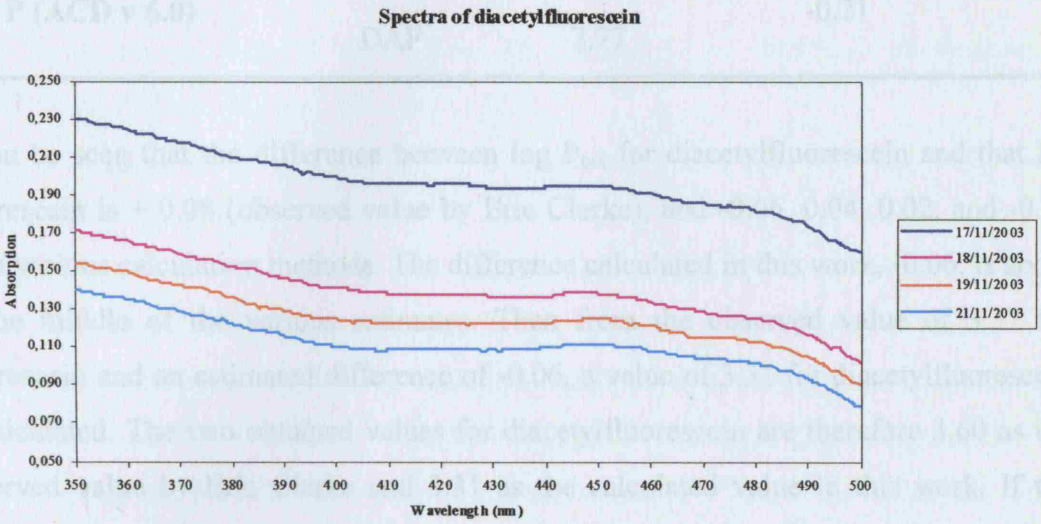


Figure 9.8. Spectra of diacetylfluorescein in a cuvette with a path length of 5 cm.

As can be observed in the spectra the structure of diacetylfluorescein seems to change over a few days, so it is impossible to carry on with the experiment.

One of the alternatives to obtain the log P value for diacetylfluorescein could be by comparing of the value of log P_{oct} obtained for fluorescein in this work and the log P_{oct} value kindly obtained by Eric Clarke for diacetylfluorescein using an HPLC method. In Table 9.13 are details.

Table 9.13
Comparison of calculated values for diacetylfluorescein with the observed value by Eric Clarke. F = Fluorescein. DAF = Diacetylfluorescein.

References	Solute	Log P	DAF – F
Eric Clarke	F	3.52 (obs)	0.08
	DAF	3.60 (obs)	
Log P_{obs} This work	F	3.37 (obs)	-0.06
Log P_{cal} This work	DAF	3.31	
Clog P (v 4.0)	F	2.70	0.04
	DAF	2.74	

Table 9.13 *(Continued)*

References	Solute	Log P	DAF – F
Clog P (v 2.0)	F	2.84	0.02
	DAF	2.86	
Log P (ACD v 6.0)	F	2.98	-0.21
	DAF	2.77	

It can be seen that the difference between log P_{oct} for diacetylfluorescein and that for fluorescein is + 0.08 (observed value by Eric Clarke), and -0.06, 0.04, 0.02, and -0.21 from various calculation methods. The difference calculated in this work, -0.06, is about in the middle of the various estimates. Then from the observed value of 3.37 for fluorescein and an estimated difference of -0.06, a value of 3.31 for diacetylfluorescein is calculated. The two obtained values for diacetylfluorescein are therefore 3.60 as the observed value by Eric Clarke and 3.31 as the calculated value in this work. If the observed value of Eric Clarke for fluorescein is too high by 0.15 log unit (3.52 – 3.37) it might be that the same applies to his value for diacetylfluorescein, and a correction to his observed value (3.60 – 0.15) would give 3.45 as the value for diacetylfluorescein. By chance, this is the average of the obtained values of 3.60 and 3.31

The descriptors in Table 9.9 for diacetylfluorescein yield, of course, 3.31 for log P_{oct}. A slightly amended set, **E** = 2.46, **S** = 3.32, **A** = 0.00, **B** = 1.46 and **V** = 2.8421 yields log P_{oct} = 3.45, and it is suggested that this amended set be used to calculate other log P values for diacetylfluorescein. The Solver calculation also yields **L** = 15.55 and log K_W = 14.33 units.

9.8. CONCLUSION

There are considerable problems in dealing with diacetylfluorescein. Firstly, the solubility in water is very small, which makes the measurement of partition coefficients difficult. Secondly, and most importantly, the spectrum of diacetylfluorescein alters with time, over a period of hours. This means that the shake flask method cannot be used to measure partition by the UV/Visible analytical method. The HPLC method of Eric Clarke is a very rapid procedure, hence Clarke was able to obtain what seems to be a reliable value of log P_{oct} for diacetylfluorescein. Using the experimental value of Eric Clarke, and taking into account the various calculations, a value of 3.45 is suggested as

log P_{oct} for diacetylfluorescein. The corresponding descriptors for diacetylfluorescein are **E** = 2.46, **S** = 3.32, **A** = 0.00, **B** = 1.46 and **V** = 2.8421.

9.9. REFERENCES

1. J. B. M. M. van Bree, A. G. de Boer, M. Danhof, L. A. Ginsel, D. D. Bremier, J. Pharmacol. Expt. Therap. 247 (1988) 1233
2. M. J. Kamlet, H. E. Ungnade (eds.), *Organic electronic spectral data*, 1946-1967, Interscience, London, 1946-1973
3. R. Sjöback, J. Nygren, M. Kubista, Spectrochimica Acta Part A 51 (1995) L7
4. W. G. David, (1999) *Pharmaceutical Analysis*, Churchill Livingston, London
5. SPARC On-line Calculator Available: http://ibmlc2.chem.uga.edu/sparc/smiles/smiles.cfm?CFID=97902&CFTOKEN=55621247 (Accessed: March 13, 2007)
6. T.W. Graham Solomons, (1988) *Organic chemistry*, 4th ed., John Wiley & Sons, New York

Chapter 10 Characterization of GLC stationary phases by Abraham's method.

10.0. INTRODUCTION

The main purpose of this task is to characterise the GLC stationary phase of a gas chromatographic column by means of the Abraham Solvation Equation in the form of Eq. (10.1), see Chapter 4. Characterising a column is a useful way to quantify the selectivity differences between stationary phases [1].

$$SP = c + e \cdot \mathbf{E} + s \cdot \mathbf{S} + a \cdot \mathbf{A} + b \cdot \mathbf{B} + l \cdot \mathbf{L} \qquad (10.1)$$

In the present case, the dependent variable, SP, is the logarithm of the relative retention time, log t_{rel}, of a series of solutes on a given GLC phase at a given temperature. Since the descriptors are supposed to represent the solute effect on various solute-phase interactions, the system constants in the model (Eq. 10.1) will convey information on the ability of the stationary phases to participate in solute-solvent intermolecular interactions, as mentioned in section 5.4.3 of this work. The regression coefficients are thus very important, because they will encode stationary phases properties. The *e* constant refers to the ability of the stationary phase to interact with solute n- or π–electron pairs, the *s* constant establishes the ability of the stationary phase to take part in dipole-type interactions, the *a* constant is a measure of stationary phase hydrogen-bond basicity (because a basic phase will interact with an acidic solute), and the *b* constant is a measure of stationary phase hydrogen-bond acidity. The *l* constant incorporates contributions from stationary phase cavity formation and solute-solvent dispersion interactions, and more specifically in gas chromatography indicates how well the phase will separate members of a homologous series [2].

The Abraham Solvation Equation, which is based on the solvation parameter model as mentioned in Chapter 4, has been applied successfully to the characterisation of a very large number of GLC stationary phases [3,4,5,6].

The objective is to obtain the coefficients *c*, *e*, *s*, *a*, *b* and *l* of Eq. (10.1) for different stationary phases by standard procedures involving multiple linear regression

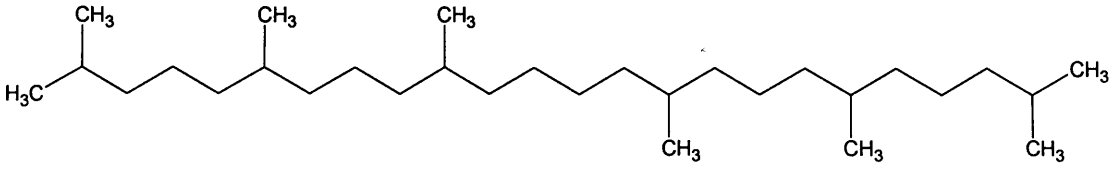
analysis (MLRA). It is obvious that not all the coefficients have to be present in the equation for each column; this depends on the selective properties of the column and, of course, some of the coefficients can be zero.

10.1. EXPERIMENTAL

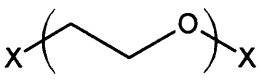
10.1.1 Materials

The stationary phases characterized in this work are squalane, carbowax 20M and diethylene glycol succinate, DEGS, whose structures are represented in Figure 10.1. Squalane was obtained from May & Baker Ltd., Dagenham, England. Carbowax 20M and DEGS were purchased from Thames Restek UK Ltd., Saunderton, England.

a. Squalane



b. Carbowax 20M



c. Diethylene glycol succinate

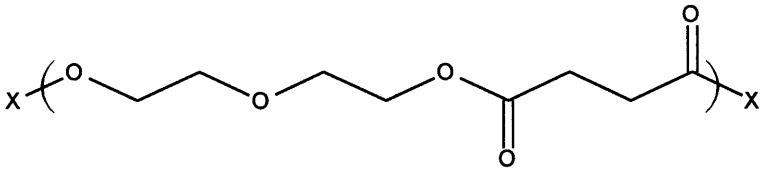


Figure 10.1. Structures of squalane, carbowax 20M and diethylene glycol succinate.

The squalane column was prepared in the laboratory. The support (Chromosorb G-AW, DMCS, 45/60 mesh) used to prepare this column was obtained from Alltech Associates Applied Science Ltd., Carnforth, England. The other two stationary phases, i.e. carbowax 20M and DEGS, were purchased already with the support (Chromosorb-G-AW-DMCS 45/60 mesh), ready to be packed into the glass column, also from Alltech Associates Applied Science Ltd., Carnforth, England.

The probe solutes are listed in the appendix of this chapter in Table 10.11, 10.12 and 10.13, along with relevant solvation parameters. All materials were 95-99% purity. They were selected from a long list of compounds for which solvation descriptors were known and were available in the laboratory of UCL where this work was carried out. Several factors were considered in selecting the final solutes. First, they represent a wide variety of compounds and functional groups. Second, the values of the solvation parameters cover a wide range of known values for a given parameter. Finally, the retention times of the solutes were not impracticably long under the experimental conditions as to make the measured retention times unreliable.

The stationary phases were selected based on the highest regression coefficients for each one. For instance, squalane was selected because of the high *l* value and the low *s*, *a* and *b* values. This phase has been previously characterised by Linear Free Energy Relationship (LFER) methods [5,6]. It was characterised with the purpose of being used in future work to calculate **L** values for compounds. Carbowax 20M was selected because of the high *e* value, and DEGS because of the high *s* and *a* value. Thus, carbowax 20M was characterised with the purpose of being used in future work to calculate **E** values, and DEGS to calculate **S** and **A** values.

10.1.2 Column preparation

Squalane

Column packaging was prepared by dissolving approximately 21 g of squalane in 150 ml of hexane in an evaporating dish, and adding approximately 142 g of support to create a slurry. The solvent was slowly evaporated, with gentle stirring from time to time, to leave the stationary phase coated on the support. Coated support was then packed into a glass column. The packed column was conditioned for 24 h at 100°C prior to performing retention studies.

Carbowax 20M and DEGS

It was only necessary to pack these two stationary phases into the glass columns as they were purchased with the support. The packed columns were then conditioned for 24 h at 100°C prior to performing retention studies, as was done for squalane.

10.1.3 Instrumentation

Squalane and carbowax 20M

Gas chromatographic measurements were made using two Perkin-Elmer gas chromatographs (model F33), one for each column, fitted with a flame ionization detector and a heated on-column injector.

DEGS

Gas chromatographic measurements were made using a PYE UNICAM gas chromatograph (model PU 4550) fitted with a flame ionization detector and a heated on-column injector.

A thermometer, type no. TC 1100, supplied by Tempcon Instrumentation Ltd. was used to measure ambient and column temperatures.

Multiple linear regression analysis was performed on a personal computer (University College London, London, UK) using Minitab, a statistical software purchased from Minitab Ltd., Coventry, England. The explanatory variables used in the data analysis are summarized in the appendix of this chapter in tables 10.11, 10.12, and 10.13, and were taken from the extensive Abraham database available at University College London, UCL.

10.1.4 Retention studies

As mentioned, the retention studies for these three columns, i.e. squalane, carbowax 20M and DEGS, were performed on a Perkin-Elmer gas chromatograph (model F33) for the first two, and on a PYE UNICAM gas chromatograph (model PU 4550) for the third one, all fitted with a flame ionization detector and a heated on-column injector. Nitrogen carrier pressure was altered over a wide range to provide a column flow-rate between 5-85 ml · min^{-1}. The column temperature was maintained at 100 °C in the cases of squalane and carbowax 20M, and at 90 °C in the case of DEGS. All quoted retention times have been corrected for column hold-up.

Squalane

Solute retention times were used to calculate the logarithm relative retention time, i.e. normalize to octane, using the following relationship

$$\log t_{rel} = \frac{\log t_R}{\log t_{R(oct)}} \qquad (10.2)$$

where t_R is the solute retention time and $t_{R(oct)}$ is the retention time of octane both taken in the same run.

In some cases this was impossible to perform because some of the compounds used in this study have a retention time far from the retention time of octane. In these cases, it was necessary to change the flow-rate of the carrier gas to a very high level and this made it impossible to obtain the retention time for octane. In these cases, an average of the relative retention time to octane of another n-alkane (the nearest in retention time to the compound in question), was taken as a reference to calculate what would be the retention time of octane. Once this retention time was calculated it was used in the same way as explained above.

Carbowax 20M

As for squalane, solute retention times were used to calculate the logarithm relative retention time, but in this case normalized to heptanol as the standard, using the following relationship

$$\log t_{rel} = \frac{\log t_R}{\log t_{R(hepOH)}} \qquad (10.3)$$

where t_R is the solute retention time and $t_{R(hepOH)}$ is the retention time of heptanol both taken in the same run.

In cases where heptanol could not be used as a standard, another alcohol was selected, and a retention time obtained for heptanol in the same way as detailed for squalane.

DEGS

The procedure is same as with carbowax 20M.

Normalizing the retention time of all the compounds to one compound chosen as a reference, reduces errors due to small changes in the flow-rate of the carrier gas during

a test, and small changes in temperature, because they will not affect the relative retention times.

10.2. RESULTS AND CONCLUSIONS

10.2.1 Multiple linear regression analysis

The retention data experimentally determined for the solutes are summarized in Table 10.1, 10.2 and 10.3, as t_{rel} and log t_{rel} values. These data were used as the dependent variable in a LFER of the form of Eq. (10.4) to obtain the LFER coefficients (*c*, *e*, *s*, *a*, *b*, *l*) for the stationary phases by MLR.

$$\log t_{rel} = c + e \cdot \mathbf{E} + s \cdot \mathbf{S} + a \cdot \mathbf{A} + b \cdot \mathbf{B} + l \cdot \mathbf{L} \qquad (10.4)$$

Table 10.1
Summary of log t_{rel} values on squalane.

Solute	t_{rel}	log t_{rel}	Solute	t_{rel}	log t_{rel}
2-propen-1-ol	0.115	-0.939	nonane	2.062	0.314
pentene	0.136	-0.866	benzonitrile	2.126	0.328
nitroethane	0.188	-0.726	heptan-1-ol	2.132	0.329
2-methylpropan-1-ol	0.200	-0.699	diiodomethane	2.177	0.338
isopropyl acetate	0.245	-0.611	octan-2-one	2.643	0.422
hexane	0.250	-0.602	phenetole	3.044	0.483
1,2-dimethoxyethane	0.257	-0.590	octan-2-ol	3.088	0.490
2-methylbutan-2-ol	0.258	-0.588	diethyl malonate	3.146	0.498
2,4-dimethylpentane	0.308	-0.511	benzyl alcohol	3.178	0.502
2-ethoxyethanol	0.347	-0.460	2-ethylhexan-1-ol	3.372	0.528
dipropylether	0.378	-0.423	decane	4.269	0.630
heptane	0.492	-0.308	5-nonanone	5.071	0.705
N,N-dimethylformamide	0.581	-0.236	cyclooctanol	8.397	0.924
2-methoxyethyl acetate	0.690	-0.161	undecane	8.812	0.945
butyl acetate	0.704	-0.152	1,3-dimethoxybenzene	8.992	0.954
toluene	0.742	-0.130	propiophenone	9.117	0.960
octane	1.000	0.000	1,2,3,4-tetramethylbenzene	11.736	1.070

Table 10.1 *(continued)*

Solute	t_{rel}	log t_{rel}	Solute	t_{rel}	log t_{rel}
N,N-dimethylacetamide	1.064	0.027	dodecane	18.169	1.259
heptan-2-one	1.286	0.109	dimethyl phthalate	41.955	1.623
cyclohexanol	1.302	0.115	tetradecane	77.263	1.888
anisole	1.809	0.257			

Table 10.2
Summary of log t_{rel} values on carbowax 20M.

Solute	t_{rel}	log t_{rel}	Solute	t_{rel}	log t_{rel}
dipropylether	0.031	-1.509	hexanol	0.568	-0.246
ethanol	0.061	-1.215	tetradecane	0.707	-0.151
benzene	0.075	-1.125	ethoxybenzene	0.729	-0.137
propyl acetate	0.082	-1.086	pentachloroethane	0.912	-0.040
isobutyl acetate	0.098	-1.009	heptanol	1.000	0.000
propanol	0.102	-0.991	hexachloro-1,3-butadiene	1.292	0.111
butyl acetate	0.135	-0.870	octanol	1.753	0.244
undecane	0.143	-0.845	hexadecane	2.053	0.312
butanol	0.182	-0.740	2-undecanone	2.194	0.341
ehtylbenzene	0.195	-0.710	ethan-1,2-diol	2.570	0.410
butyl propanoate	0.197	-0.706	o-tolunitrile	2.792	0.446
4-methylpentan-2-ol	0.198	-0.703	nonanol	3.038	0.483
1-bromohexane	0.216	-0.666	triethyl phosphate	3.257	0.513
dodecane	0.245	-0.611	heptadecane	3.420	0.534
limonene	0.282	-0.550	p-dibromobenzene	3.742	0.573
propylbenzene	0.299	-0.524	1,2,3-trichlorobenzene	3.838	0.584
pentanol	0.325	-0.488	nitrobenzene	4.099	0.613
tert-butylbenzene	0.342	-0.466	o-nitrotoluene	5.107	0.708
2-methoxyethyl acetate	0.348	-0.458	decanol	5.309	0.725
γ-terpinene	0.352	-0.453	formamide	6.434	0.808
hexyl acetate	0.392	-0.407	m-chloronitrobenzene	8.821	0.946
2-ethoxyethyl acetate	0.428	-0.369	undecanol	9.055	0.957
2-chlorotoluene	0.490	-0.310	dodecanol	15.721	1.196

Table 10.3
Summary of log t_{rel} values on DEGS.

Solute	t_{rel}	log t_{rel}	Solute	t_{rel}	log t_{rel}
2,2,4-trimethylpentane	0.019	-1.721	1,2,3,4-tetramethylbenzene	1.294	0.112
heptane	0.020	-1.699	heptadecane	1.393	0.144
ethanol	0.085	-1.071	octanol	1.644	0.216
4-methylpentan-2-one	0.129	-0.889	N,N-dimethylformamide	1.730	0.238
propanol	0.131	-0.883	linalool	1.914	0.282
1,2-dichloroethane	0.147	-0.833	undecan-2-one	2.084	0.319
butyl propanoate	0.209	-0.680	formic acid	2.286	0.359
1-heptaldehyde	0.283	-0.548	diethylformamide	2.661	0.425
tetradecane	0.299	-0.524	nonanol	2.704	0.432
nitromethane	0.315	-0.502	N,N-dimethylacetamide	3.112	0.493
nitroethane	0.336	-0.474	benzonitrile	3.170	0.501
pentanol	0.366	-0.437	1,2,3-trichlorobenzene	3.758	0.575
butylbenzene	0.449	-0.348	p-dibromobenzene	3.776	0.577
octan-2-one	0.488	-0.312	o-tolunitrile	3.811	0.581
2-chlorotoluene	0.495	-0.305	decanol	4.436	0.647
1-dodecyne	0.533	-0.273	naphthalene	5.140	0.711
nonan-2-one	0.793	-0.101	nitrobenzene	6.266	0.797
ethoxybenzene	0.910	-0.041	undecanol	7.228	0.859
heptanol	1.000	0.000	o-nitrotoluene	7.638	0.883
diiodomethane	1.213	0.084	dodecanol	11.641	1.066
1,2-dichlorobenzene	1.256	0.099			

10.2.2 Results

LFER results for each stationary phase used in this work are summarized in Table 10.4.

Table 10.4
Linear Free Energy Relationship (LFER) coefficients[a].

Stationary phase	*c*	*e*	*s*	*a*	*b*	*l*
squalane	-2.23	0.15	-0.03	-0.01	0.05	0.61
	(0.01)	(0.02)	(0.02)	(0.04)	(0.03)	(0.00)
carbowax 20M	-3.36	0.32	1.23	1.76	0.33	0.48
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.00)
DEGS	-3.06	0.26	1.54	2.06	0.00	0.39
	(0.06)	(0.05)	(0.07)	(0.09)	(0.09)	(0.01)

[a] Standard error associated with individual coefficients provided in parentheses.

Squalane

Four compounds of the forty-five used to obtain the coefficients of Eq. 10.4 were found to be outliers in the regression analysis, and were consequently left out of the set used to obtain the final coefficients. These compounds were 5-methylheptan-3-one, propan-2-ol, ethan-1,2-diol, and hexachloro-1,3-butadiene.

Previous LFER characterizations were performed for squalane by several groups using chromatographic retention data reported by Abraham et al. [6] and Poole and Kollie [5]. Results of those studies are included in Table 10.5 for comparison.

There are some differences between the coefficients systems values reported in Table 10.5. The coefficient values in the first line were obtained using retention data, log K values, reported by Poole and co-workers [7] in a previous work. The coefficient values in the second line were reported by Abraham et al. and can also be found in the reference indicated in the table. The temperature at which these retention data were obtained is 121.4°C. Thus, these coefficient values cannot be directly compared with the results of this work. These data are included here because they seem to be amongst the most reliable GLC data ever reported, with considerable care being taken to exclude contributions from interfacial adsorption [6].

Table 10.5
Comparison of LFER coefficients for squalane.

	c	*e*	*s*	*a*	*b*	*l*	T (°C)	Ref
Poole and Kollie	-2.10	0.11		-0.06	0.03	0.64	101.3	[5]
Abraham et al.	-0.20	0.13	0.02	-0.10		0.58	121.4	[6]
This work	-2.23	0.15	-0.03	-0.01	0.05	0.61	99.8	

The coefficient system of Poole and Kollie in Table 10.5 can be found in the references indicated in the table and were obtained using retention data reported in Ref. [6,8-10]. The regression statistics found in these two works are also included in Table 10.6, together with those found in this work, so that the reader can compare them. The fact that the result of this work is in a good agreement with previously reported values suggests that the results presented here appear quite reasonable.

It should be noted that the *s*, *a* and *b* coefficient systems were found to be not statistically significant as judged by the *t-test* in this study. Besides, the characteristic phase constants cannot be negative except for *e* in a few phases that contain fluorine

atoms [1], and *c* for any phase. Thus, the best statistical fit may not correspond to the fit that makes chemical sense. The *s* and *b* values were found also to be not statistically significant in previous Poole [5] and Abraham [6] works. So the exclusion of this two terms *s* and *b*, together with the term *a* lead to Eq. (10.5), where the correlation coefficient, R^2, and the *F-statistic*, *F*, still remain satisfactory.

Table 10.6
Regression Statistics for Eq. 10.4 for squalane.

	N	R^2	SD	F	Ref
Poole and Kollie	62	1.000	0.011		[5]
Abraham et al.	39	0.996	0.033		[6]
This work	41	0.999	0.022	7451	

$$\log t_{rel} = -2.22 + 0.14\,\mathbf{E} + 0.61\,\mathbf{L} \qquad (10.5)$$

N= 41; SD = 0.023; R^2 = 0.999 ; F = 17461

The success in the application of this truncated solvation equation (Eq. 10.5) to non-polar gas-liquid chromatography stationary phases has been described previously by Abraham and co-workers [8,9,11].

Carbowax 20M

Four compounds of the fifty used to obtain the coefficients of Eq. 10.4 were found outliers in the regression analysis, and consequently left out of the set used to obtain the final coefficients, viz: N,N-dimethylacetamide, diethyl malonate, propan-1,2-diol, and octanoic acid.

As for squalane, previous LFER characterizations were also performed for carbowax 20M by Poole and Kollie [5], and Abraham et al. [6]. Results of those studies are summarized in Table 10.7 for comparison.

Table 10.7
Comparison of LFER coefficients for carbowax 20M

	c	*e*	*s*	*a*	*b*	*l*	T (°C)	Ref
Poole and Kollie	-0.48	0.25	1.44	2.06	-0.18	0.48	101.3	[5]
Abraham et al.	-0.56	0.29	1.29	1.80		0.45	121.4	[6]
This work	-3.36	0.32	1.23	1.76	0.33	0.48	101.4	

Some differences between the coefficients systems values reported in Table 10.7 are found for this stationary phase. However, the coefficients system values reported in the three different studies are all of the same order, a fact that makes the results presented here reasonable. The regression statistics found previously are also included in Table 10.8, together with those found in this work, for comparison.

Table 10.8
Regression Statistics for Eq. 10.4 for carbowax 20M.

	N	R^2	SD	F	Ref
Poole and Kollie	62	0.999	0.013		[5]
Abraham et al.	39	0.997	0.059		[6]
This work	46	0.997	0.036	3169	

It should be noted that the *b* coefficient system was found to be statistically significant as judged by the *t-test*. This is unexpected for this stationary phase. As can be seen from Figure 10.1, this stationary phase appears to lack of hydrogen-bond donor groups. However, it is quite possible that the phase may have some acidity (impurity) in the stationary phase and that is why the *b* coefficient is statistically significant in the regression analysis.

Diethylene glycol succinate

One compound of the forty-two used to obtain the coefficients of Eq. 10.4 was found to be an outlier in the regression analysis, and consequently was left out of the set used to obtain the final coefficients, i.e. 2,2,2-trifluoroethanol.

As for squalane and carbowax 20M, previous LFER characterizations were also performed for carbowax 20M by Poole and Kollie [5], and Abraham et al. [6]. Results of those studies are summarized in Table 10.9 for comparison.

Table 10.9
Comparison of LFER coefficients for DEGS.

	c	*e*	*s*	*a*	*b*	*l*	T (°C)	Ref
Poole and Kollie	-0.65	0.23	1.56	2.10	0.19	0.41	101.3	[5]
Abraham et al.	-0.50	0.35	1.68	1.72		0.31	121.4	[6]
This work	-3.06	0.26	1.54	2.06	0.00	0.39	90.3	

Observing this table, it is clear that the two studies carried out at the same temperature (i.e. Poole and Kolie, and this work) yield coefficients that are very similar, and are a little way away from the coefficients found by Abraham and co-workers at 121.4 °C. The regression statistics found in previous studies are also included in Table 10.10, together with the regression statistics found in this work.

Table 10.10
Regression Statistics for Eq. 10.4 for DEGS.

	N	R^2	SD	F	Ref
Poole and Kollie	62	0.999	0.015		[5]
Abraham et al.	38	0.989	0.096		[6]
This work	41	0.988	0.077	726	

It should be noted that the *b* system constant was found to be not statistically significant as judged by the *t-test* in this study. And the exclusion of this term leads to Eq (10.6), where the correlation coefficient, R^2, and the *F-statistic*, *F*, still remain satisfactory.

$$\log tr = -3.06 + 0.26\,\mathbf{E} + 1.54\,\mathbf{S} + 2.06\,\mathbf{A} + 0.39\,\mathbf{L} \qquad (10.6)$$

$N = 41$; $SD = 0.077$; $R^2 = 0.988$; $F = 726$

10.2.3 Interpretation of LFER coefficient values

Squalane

As expected from the structure of squalane, Figure 10.1, the *a* and *b* values of Eq. 10.4 are not significant for this stationary phase, since this stationary phase lacks hydrogen bond acceptors and donors. This coating exhibits a negative *s* value, which indicates that the coating has no ability to take part in dipolarity-polarizabilty interactions. Finally, squalane does exhibit relatively large values of *e* and *l*, indicating the tendency of the phase to interact with solutes through π- and n-electron pairs, and the hydrophobicity of the phase, respectively.

Carbowax 20M

As mentioned above, the *b* coefficient system was found to be statistically significant, something unexpected for this stationary phase taking into account the structure of

carbowax 20M. This does not mean that the structure of carbowax 20M has hydrogen bond donors, but some acid impurity may be present in the stationary phase. On the other hand, there are present hydrogen bond acceptors in this structure, leading to the high value of the *a* system constant, which represents high number of hydrogen bond donor interactions. The largest value of *e*, comparing with the other two stationary phases, indicates the large tendency of the coating to interact with solutes through π- and n-electron pairs. The *l* value is also relatively large, indicating the hydrophobicity of the phase.

Diethylene glycol succinate

As expected from the structure of DEGS, Figure 10.1, the *b* value of Eq 10.4 is not significant for this stationary phase, since this stationary phase lacks hydrogen bond donors. However, this stationary phase can interact with solutes through hydrogen bond acceptor interactions (see Figure 10.1), leading to the largest value of the *a* system constant of the three coatings studied here. This coating also exhibits a positive and relatively large *e*, which indicates the tendency of the phase to interact with solutes through π- and n-electron pairs. The largest value of *s*, compared with the other two stationary phases, indicates the large ability of this coating to take part in dipolarity-polarizability interactions. The *l* value is also considerably large, indicating the hydrophobicity of the phase.

10.3. REFERENCES

1. M. H. Abraham, C. F. Poole, S. K. Poole, J Chromatogr. A 842 (1999) 79
2. C. F., Poole, *The Essence of Chromatography*, Elsevier Science B.V, Amsterdam, 2003
3. M. H. Abraham, Anal. Chem. 69 (1997) 613
4. W. Ti, D. S. Ballantine, Jr., J Chromatogr. A 718 (1995) 357
5. C. F., Poole, T. O. Kollie, Anal. Chim. Acta 282 (1993) 1
6. M. H. Abraham, G. S. Whiting, R. M. Doherty, W. J. Shuely, J. Chromatogr. 587 (1991) 229
7. B. R. Kersten, S. K. Poole, C. F. Poole, J. Chromatogr. 468 (1990) 329
8. M. H. Abraham, G. S. Whiting, R. M. Doherty, W. J. Shuely, J. Chromatogr. 587 (1991) 213

9. M. H. Abraham, G. S. Whiting, J. Chromatogr. 594 (1992) 229
10. M. H. Abraham, G. S. Whiting, R. M. Doherty, W. J. Shuely, P. Sakellariou, Polymer 33 (1992) 2162
11. M. H. Abraham, J. Chromatogr. 644 (1993) 95

10.4. APPENDIX

Table 10.11
Solutes and descriptors used in Eq. 10.1 for squalane

Solute	E	S	A	B	L
propan-2-ol	0.121	0.36	0.33	0.56	1.764
2-propen-1-ol	0.342	0.46	0.38	0.48	1.951
pentene	0.000	0.00	0.00	0.00	2.162
2-methylpropan-1-ol	0.217	0.39	0.37	0.48	2.413
isopropyl acetate	0.550	0.57	0.00	0.47	2.546
2-methylbutan-2-ol	0.194	0.30	0.31	0.60	2.630
1,2-dimethoxyethane	0.116	0.67	0.00	0.68	2.654
ethan-1,2-diol	0.404	0.90	0.58	0.78	2.661
hexane	0.000	0.00	0.00	0.00	2.668
2-ethoxyethanol	0.237	0.52	0.31	0.81	2.792
2,4-dimethylpentane	0.000	0.00	0.00	0.00	2.809
dipropylether	0.008	0.25	0.00	0.45	2.954
heptane	0.000	0.00	0.00	0.00	3.173
2-methoxyethyl acetate	0.166	0.79	0.00	0.81	3.290
toluene	0.601	0.52	0.00	0.14	3.325
butyl acetate	0.071	0.60	0.00	0.45	3.353
N,N-Dimethylacetamide	0.363	1.35	0.00	0.77	3.638
octane	0.000	0.00	0.00	0.00	3.677
cyclohexanol	0.460	0.54	0.32	0.57	3.758
heptan-2-one	0.123	0.68	0.00	0.51	3.760
anisole	0.708	0.75	0.00	0.29	3.890
benzonitrile	0.742	1.11	0.00	0.33	4.039
heptan-1-ol	0.211	0.42	0.37	0.48	4.115
nonane	0.000	0.00	0.00	0.00	4.182
5-methylheptan-3-one	0.110	0.63	0.00	0.51	4.200

Table 10.11 *(continued)*

Solute	E	S	A	B	L
benzyl alcohol	0.803	0.87	0.39	0.56	4.221
phenetole	0.681	0.70	0.00	0.32	4.242
octan-2-one	0.108	0.68	0.00	0.51	4.257
octan-2-ol	0.158	0.36	0.33	0.56	4.339
2-ethylhexan-1-ol	0.209	0.39	0.37	0.48	4.433
decane	0.000	0.00	0.00	0.00	4.686
5-nonanone	0.103	0.66	0.00	0.51	4.698
1,3-dimethoxybenzene	0.816	1.01	0.00	0.45	5.022
cyclooctanol	0.566	0.54	0.32	0.58	5.054
1,2,3,4-tetramethylbenzene	0.794	0.66	0.00	0.19	5.176
propiophenone	0.804	0.95	0.00	0.51	4.971
dodecane	0.000	0.00	0.00	0.00	5.696
dimethyl phthalate	0.780	1.40	0.00	0.84	6.051
tetradecane	0.000	0.00	0.00	0.00	6.705

Table 10.12
Solutes and descriptors used in Eq. 10.1 for carbowax 20M

Solute	E	S	A	B	L
dipropylether	0.008	0.25	0.00	0.45	2.954
ethanol	0.246	0.42	0.37	0.48	1.485
benzene	0.610	0.52	0.00	0.14	2.786
propyl acetate	0.092	0.60	0.00	0.45	2.819
isobutyl acetate	0.052	0.57	0.00	0.47	3.161
propanol	0.236	0.42	0.37	0.48	2.031
butyl acetate	0.071	0.60	0.00	0.45	3.353
undecane	0.000	0.00	0.00	0.00	5.191
butanol	0.224	0.42	0.37	0.48	2.601
ehtylbenzene	0.613	0.51	0.00	0.15	3.778
butyl propanoate	0.058	0.56	0.00	0.47	3.833
4-methyl-2-pentanol	0.167	0.33	0.33	0.56	3.179
1-bromohexane	0.349	0.40	0.00	0.12	4.130
dodecane	0.000	0.00	0.00	0.00	5.696
limonene	0.488	0.28	0.00	0.21	4.725

Table 10.12 *(continued)*

Solute	E	S	A	B	L
propylbenzene	0.604	0.50	0.00	0.15	4.230
pentanol	0.219	0.42	0.37	0.48	3.106
tert-butylbenzene	0.619	0.49	0.00	0.18	4.413
γ-terpinene	0.497	0.32	0.00	0.20	4.815
hexyl acetate	0.056	0.60	0.00	0.45	4.290
2-ethoxyethyl acetate	0.099	0.79	0.00	0.79	3.747
2-chlorotoluene	0.762	0.65	0.00	0.07	4.173
hexanol	0.210	0.42	0.37	0.48	3.610
tetradecane	0.000	0.00	0.00	0.00	6.705
phenetole	0.681	0.70	0.00	0.32	4.242
heptanol	0.211	0.42	0.37	0.48	4.115
hexachloro-1,3-butadiene	1.019	0.52	0.00	0.10	5.421
octanol	0.199	0.42	0.37	0.48	4.619
hexadecane	0.000	0.00	0.00	0.00	7.714
2-undecanone	0.101	0.68	0.00	0.51	5.732
ethan-1,2-diol	0.404	0.90	0.58	0.78	2.661
o-tolunitrile	0.780	1.06	0.00	0.31	4.478
nonanol	0.193	0.42	0.37	0.48	5.120
heptadecane	0.000	0.00	0.00	0.00	8.218
p-dibromobenzene	1.150	0.86	0.00	0.04	5.324
1,2,3-trichlorobenzene	1.030	0.86	0.00	0.00	5.419
o-nitrotoluene	0.866	1.11	0.00	0.28	4.878
decanol	0.191	0.42	0.37	0.48	5.610
m-chloro-nitrobenzene	1.000	1.14	0.00	0.25	5.206
undecanol	0.181	0.42	0.37	0.48	6.128
dodecanol	0.175	0.42	0.37	0.48	6.620

Table 10.13
Solutes and descriptors used in Eq. 10.1 for DEGS

Solute	E	S	A	B	L
2,2,4-trimethylpentane	0.000	0.00	0.00	0.00	3.106
heptane	0.000	0.00	0.00	0.00	3.173

Table 10.13 *(continued)*

Solute	E	S	A	B	L
ethanol	0.246	0.42	0.37	0.48	1.485
4-methylpentan-2-one	0.111	0.65	0.00	0.51	3.089
propanol	0.236	0.42	0.37	0.48	2.031
1,2-dichloroethane	0.416	0.64	0.10	0.11	2.573
2,2,2-trifluoroethanol	0.015	0.60	0.57	0.25	1.224
butyl propanoate	0.058	0.56	0.00	0.47	3.833
1-heptaldehyde	0.140	0.65	0.00	0.45	3.865
tetradecane	0.000	0.00	0.00	0.00	6.705
nitromethane	0.313	0.95	0.06	0.31	1.892
nitroethane	0.270	0.95	0.02	0.33	2.414
pentanol	0.219	0.42	0.37	0.48	3.106
butylbenzene	0.600	0.51	0.00	0.15	4.730
2-octanone	0.108	0.68	0.00	0.51	4.257
2-chlorotoluene	0.762	0.65	0.00	0.07	4.173
1-dodecyne	0.133	0.22	0.09	0.10	5.657
2-nonanone	0.113	0.68	0.00	0.51	4.735
phenetole (ethoxybenzene)	0.681	0.70	0.00	0.32	4.242
heptanol	0.211	0.42	0.37	0.48	4.115
diiodomethane	1.453	0.69	0.05	0.23	3.857
1,2-dichlorobenzene	0.872	0.78	0.00	0.04	4.518
1,2,3,4-tetramethylbenzene	0.794	0.66	0.00	0.19	5.176
heptadecane	0.000	0.00	0.00	0.00	8.218
octanol	0.199	0.42	0.37	0.48	4.619
n,n-dimethylformamide	0.367	1.31	0.00	0.74	3.173
linalool	0.398	0.55	0.20	0.67	4.794
2-undecanone	0.101	0.68	0.00	0.51	5.732
formic acid	0.343	0.75	0.76	0.33	1.545
diethylformamide	0.305	1.25	0.00	0.76	3.995
nonanol	0.193	0.42	0.37	0.48	5.120
n,n-dimethylacetamide	0.363	1.35	0.00	0.77	3.638
benzonitrile	0.742	1.11	0.00	0.33	4.039
1,2,3-trichlorobenzene	1.030	0.86	0.00	0.00	5.419

Table 10.13 *(continued)*

Solute	E	S	A	B	L
p-dibromobenzene	1.150	0.86	0.00	0.04	5.324
o-tolunitrile	0.780	1.06	0.00	0.31	4.478
decanol	0.191	0.42	0.37	0.48	5.610
naphthalene	1.340	0.92	0.00	0.20	5.161
nitrobenzene	0.871	1.11	0.00	0.28	4.557
undecanol	0.181	0.42	0.37	0.48	6.128
o-nitrotoluene	0.866	1.11	0.00	0.28	4.878
dodecanol	0.175	0.42	0.37	0.48	6.620

Chapter 11 Determination of Solvation Properties of Volatile Organic Compounds from GLC Data

11.0. INTRODUCTION

The use of properties that are easy to measure in order to calculate or estimate properties that are difficult to measure is a well known method in most fields of chemistry and biochemistry. With the advance of modern techniques such as combinatorial chemistry, high throughput screening has become very important and therefore estimates of physical properties from structure by calculations that can be performed rapidly are of primary importance. A large number of transport-related processes involve either the equilibrium transfer (K or P) or the rate of transfer (k) of a solute from one phase to another. Since log K (or log P) and log k are free-energy related, Abraham and co-workers formulated a number of solute properties or descriptors that are also free-energy related and could be used for the correlation of log K and log k values.

The original work of Kamlet and Taft and co-workers showed that it was indeed possible to define a small number of descriptors that could be combined in a linear way for the correlation of solute properties [1,2]. After considerable preliminary work [3,4], Abraham and co-workers succeeded in constructing a new and more rigorous set of five solute descriptors, [5-10] described in Chapter 4 of this work. Briefly, **E** is an excess molar refraction that is obtained from refractive index for solutes that are liquid at 20 °C. For solids, the refractive index of the hypothetical liquid at 20 °C can be calculated, or **E** can be obtained by the summation of fragments or substructures. **S** is the dipolarity/polarizability that can be obtained from gas liquid chromatographic measurements on polar stationary phases, or more generally from water–solvent partition coefficients. **A** and **B** are the overall or effective hydrogen bond acidity and basicity that are most easily obtained from water–solvent partitions, **V** is the McGowan characteristic volume, and **L** is the gas–liquid partition coefficient on hexadecane at 298 K. The dependent variable, SP, is a solute property in a given system. The range of solutes for which descriptors are currently available is now quite large, and encompasses compounds as far apart as helium, hydrogen, nitrogen, etc. on the one hand and drugs, environmental pollutants and pesticides on the other. These solute

descriptors can be combined into two linear free energy relationships Eqs. (11.1) and (11.2).

$$SP = c + e \cdot \mathbf{E} + s \cdot \mathbf{S} + a \cdot \mathbf{A} + b \cdot \mathbf{B} + l \cdot \mathbf{L} \quad (11.1)$$

$$SP = c + e \cdot \mathbf{E} + s \cdot \mathbf{S} + a \cdot \mathbf{A} + b \cdot \mathbf{B} + v \cdot \mathbf{V} \quad (11.2)$$

Eq. (11.1) was designed to deal with transfers from the gas phase to a condensed phase, and Eq. (11.2) for transfers from one condensed phase to another.

The coefficients *c*, *e*, *s*, *a*, *b*, *l* and *v* in the equations are found by standard procedures for multiple linear regression analysis. Descriptors for a large number of solutes have been obtained from experimental data; the maximum and minimum ranges of these descriptors in our database are shown in Table 11.1. The solute descriptors represent the solute influence on various solute–solvent phase interactions.

Table 11.1
Table of the available solute descriptors

Descriptor	Maximum value	Minimum value	Total
E	4.62	-1.21	5957
S	5.70	-1.80	5501
A	4.33	0.00	6137
B	5.68	0.00	5233
V	8.56	0.07	6194
L	34.48	-1.74	4349

Because of the ease of use, the availability of commercial equipment, and the reproducibility of measurements, it is not surprising that chromatographic methods have long been used to obtain properties or 'descriptors' that characterize compounds. Gas-liquid chromatography can also operate at very low compound concentration and so, with few exceptions, can yield descriptors for compounds as the simple monomeric species.

In this work, solvent-water partition coefficients, log P, available in the Med-Chem database [12] have been combined with chromatographic data to determine the descriptors for a large number of volatile organic compounds, VOCs, present in indoor air, see Table 11.6. These compounds are summarized in Table 11.12 in the appendix of this chapter. The aim is to obtain the best set of descriptors from as a large number of processes as possible. The bigger the number of data sets used the more

reliable the results. The Medicinal Chemistry database, organized by Leo, is a most valuable source of experimental data that can be used in the determination of descriptors.

11.1. METHODOLOGY. THE SYSTEM OF ABRAHAM

This system uses data both from GLC and HPLC, as well as data on water–solvent partition coefficients. However in this work only data from GLC and water–solvent partition coefficients available in the Med-Chem database [12] will be used.

11.1.1 The E and V descriptors

The method is quite straightforward. As starting point, the descriptor **E** can be calculated from the refractive index, if available. A computer program to calculate **E**, named VR, is available at University College London, London, UK. It requires as input the refractive index at 20°C, the solute molecular formula and the number of rings in the molecule. The **E** descriptor has been found to be very nearly an additive property, and so another method of obtaining **E** is through addition of fragments. The McGowan volume, **V**, can simply be calculated from atomic fragments and the number of bonds, B_n, in a molecule, all bonds being counted as one, no matter whether single, double or triple. As mentioned in Chapter 4 of this work, it is not necessary to count the number of bonds in complicated molecules because the algorithm of Abraham [13] can be used, Eq. (11.3), where N_a is the total number of atoms and R_g the number of rings. The VR program, above, calculates **V** using as input only the compound molecular formula, and the number of rings in the compound; N_a is of course obtained from the molecular formula.

$$B_n = N_a - 1 + R_g \tag{11.3}$$

11.1.2 The L, A and S descriptors

Abraham *et al.* [13] showed that **L** could be obtained by direct measurement of retention volumes by GLC, using a hexadecane stationary phase thermostatted at 25 °C.

Since then, Abraham and co-workers [14,15,16] have shown that GLC data, SP, for solutes on a non polar stationary phase, usually at 100 °C or higher, can be correlated through the simple equation,

$$SP = c + e \cdot \mathbf{E} + l \cdot \mathbf{L} \qquad (11.4)$$

Once the phase has been calibrated or 'characterized' with solutes of known SP, **E** and **L** values, then other **L** values can be obtained for further solutes for which SP and **E** are known. In this way, **L** values for hundreds of solutes with **L** values up to 7.71 (decylbenzene) were obtained. The determination of **L** through equations like Eq. (11.4) is a very simple direct experimental method.

Eq. (11.4) can be used to correlate GLC retention data for solutes on non polar phases such as squalane or DB-5MS, the former used in this work. However, if the stationary phase is polar, Eq. (11.4) will apply only to non-polar solutes such as alkanes. For polar compounds that have no additional hydrogen bond acidity, one additional descriptor is needed [14,15,16] that refers to the dipolarity/polarizability of the solute, that is descriptor **S**. Thus for relative retention times, t_{rel}, of solutes on di-n-propyl tetrachlorophthalate at 90°C, Abraham and Whiting [15] found,

$$\log t_{rel} = -3.433 + 1.640\ \mathbf{S} + 0.618\ \mathbf{L} \qquad (11.5)$$

A large number of equations of this type, or equations with the additional **E** descriptor, were used to obtain **S** descriptors, mostly for reasonably volatile solutes [14,15,16].

GLC phases that are polar are invariably also hydrogen bond bases, for example the dialkyl phthalates. Then if the solutes that are studied on such stationary phases are not hydrogen bond acids, equations such as Eq. (11.5), or equations in **E**, **S** and **L**, will suffice to correlate retention data, SP. However, if the data set includes hydrogen bond acids, then any correlation equation will require also the **A** descriptor [3,14,15,16], as shown in Eq. (11.6).

$$SP = c + e \cdot \mathbf{E} + s \cdot \mathbf{S} + a \cdot \mathbf{A} + l \cdot \mathbf{L} \qquad (11.6)$$

Then in order to obtain **A** values for compounds, not only must the stationary phase be calibrated, but values of SP, **E**, **S** and **L** must be known.

11.1.3 The B descriptor

There are very few GLC stationary phases that have been examined that are significant hydrogen bond acids, and almost no such phases are commercially available. This means that the above GLC step by step method cannot be extended to the determination of **B** values. Abraham [6] therefore turned to the use of water–solvent partition data, as log P values, in order to establish a scale of hydrogen bond basicity. The equation used was the full Eq. (11.2), where SP is now log P.

Now that a large number of such systems have been characterized through Eq. (11.2), it is possible to determine not only the descriptor **B**, but also **S** and **A**, through log P values for a given solute in a number of water systems. As mentioned above, the descriptors **E** and **V** can readily be obtained, and so only the descriptors **S**, **A** and **B** in Eq. (11.2) need to be determined.

In principle, if log P values are known for a compound in three water–solvent systems which have been characterized by Eq. (11.2), then we have three equations and three unknowns (**S**, **A** and **B**) and so the unknowns can be evaluated. In practice, this procedure will only work satisfactorily if the coefficients in the three equations are substantially different. Two stratagems can be applied:

Firstly, log P values in a large number of water–solvent systems may be known or may be determined. Then the **S**, **A** and **B** values that lead to best reproduction of the log P values can be evaluated. It is often the case that one or more log P values are out of line, and have to be left out. The second method is to use a small number of systems, but to select them carefully so that the corresponding equations are as different as possible, and therefore this will lead also to the best fit.

11.1.4 A general method for the determination of descriptors

Because there is no reason why the determination of descriptors should be restricted to GLC data or log P data, and as at this time numerous GLC and log P systems have been characterized through Eqs. (11.1) and (11.2), any combination of data can be used to determine **S**, **A**, **B** and **L**. The method is as before: **E** and **V** are known or can be calculated, and values of **S**, **A**, **B** and **L** are calculated that best reproduce the total set of experimental values, GLC data and log P data, available for the compound of interest. This is the method used in this work.

The 'Solver' facility in the Microsoft Excel spreadsheet has been used to calculate the descriptors in the way describe above. Solver is a tool in Microsoft Excel which can be used to determine the maximum or minimum value of one cell by changing other cells. Solver minimises the sum of squares on the required equations to fit the targeted cells **S**, **A**, **B** and **L** and the values are accepted when the overall sums of squares are at a minimum. Solver uses the generalised reduced gradient (GRG2) nonlinear optimisation code developed by Leon Lasdon, University of Texas at Austin, and Allan Waren, Cleveland State University.

This Solver tool is a straightforward and rapid computer program, which avoids a large number of laborious calculations. However, this is only a statistical function that does not consider the physicochemical properties of the solute. Quite often the Solver results have to be adjusted to agree with the chemical features of the compound. For instance, an '**A**' value is sometimes attributed after running Solver to a compound that does not contain any hydrogen atom. In such a case, unfitted descriptors values, e.g. $\mathbf{A} \neq 0$, are replaced by expected descriptor values, e.g. $\mathbf{A} = 0$, which are fixed in Solver, and another statistical analysis is run. A new set of descriptors is obtained with a new standard deviation that is very often close to the one originally obtained. Another example is when the output presents either **A** or **B**, or both, with a negative value. In this case the **A** or **B** value for the compound has to be fixed in Solver, or the Solver value for **A** or **B** is fixed at ≥ 0.

The above method is highly dependent upon the availability of physicochemical data for any given volatile organic compounds, VOCs. GLC data are very abundant, thus a number of equations in the line of Eq. (11.1) can be developed. Conversely, data for condensed processes, such as log P, are scarcely reported except for log P octanol. The general lack of data limits the use of equations similar to Eq. (11.2) in the task of descriptor determination. Therefore, difficulties in obtaining an experimental value for the **B** descriptor arise.

The first step in the procedure is then to construct equations of the type of Eq. (11.1) for as many stationary phases as possible, for each VOC. Data found in the literature were used to set up these equations. Furthermore, three stationary phases were characterised in the laboratory of Analytical Chemistry at University College London, London, UK, for the same purpose. On the other hand, equations of the type of Eq. (11.2) have been previously constructed by Abraham and co-workers, and a database of these equations can be found at University College London, London, UK.

11.2. RESULTS

An example in the determination of descriptors is that of the solute 1,2-dicholoropropane. This is one of the VOCs found in indoor air, see Table 11.12 in the appendix of this chapter. **E** could be obtained from the experimental liquid refractive index and the calculated value of **V**. Log P values were available in two different water-solvent systems for this compound [12], and a log K^w value of 0.937 was obtained from the vapour pressure in mm Hg, and the water solubility in mg/l, at 25 °C, via Eq. (11.7), see Table 11.2.

$$\log K^w = \log C^W - \log C^G \tag{11.7}$$

where C^W is given by,

$$C^W = \frac{\text{water solubility}}{\text{molecular weight} \cdot 1000} \tag{11.8}$$

and C^G by,

$$C^G = \frac{\text{vapour pressure}}{760\,\text{mm Hg} \cdot 24.45\,\text{L}} \tag{11.9}$$

This enables the corresponding two log K^S values to be deduced through two equations of type Eq. (11.10).

$$\log P = \log K^S - \log K^W \tag{11.10}$$

In addition, Kovats retention indices, I, were available, from the literature, in seven GLC systems. These systems can be found in the references indicated in Table 11.2. Because the absolute values of the Kovats indices are so much larger than those of log P, it was essential to weight the GLC data by dividing all the indices by a factor of 100 in order that the I equations do not have an undue proportional weight in the minimization procedure. Then these values were used to set up the corresponding

equations. If log P and Kovats indices values are used as such, then Solver minimize the differences in calculated and observed I values with little regard to the log P values.

Table 11.2
Calculation of descriptors for 1,2-dicholoropropane

System	SP	SP_{calc}	SP_{obs}	Reference
water-octanol	log P	2.003	1.990	[12]
water-hexadecane	log P	1.927	1.930	[12]
gas-water	log K^W	0.973	0.937	Eq. (11.7)
gas-octanol	log K^S	2.925	2.927	Eq. (11.10)
gas-hexadecane	log K^S	2.836	2.867	Eq. (11.10)
gas-water	log K^W	0.975	0.937	Eq. (11.7)
GLC-SPB1	I/100	6.828	6.840	[17]
GLC-DB1	I/100	6.624	7.473	[18]
GLC-DB-XLB	I/100	7.000	8.026	[18]
GLC-BP1	I/100	6.800	6.840	[19]
GLC-BP10	I/100	7.765	7.780	[19]
GLC-DB1	I/100	6.737	6.736	[20]
GLC-DB1	I/100	6.807	6.721	[20]

Knowing **E** = 0.369 and **V** = 0.7761, only three descriptors, **S**, **B** and **L**, need to be calculated because **A** = 0. With **S** = 0.63, **B** = 0.170 and **L** = 2.836, 11 of the 13 initial data were reproduced with an SD of 0.036 log units. The two systems not used are shown in italic in the above table.

A good visual way of looking at the ability of the descriptors **S**, **B** and **L**, determined for 1,2-dichloropropane to predict the solute property, SP, of this compound in the different systems in Table 11.2, is by plotting the graph of the calculated values versus the observed values for SP (Figure 11.1). The two outliers found in the analysis, in italic in Table 11.2, have been removed from the graph.

As can be observed in Figure 11.1, the points give a straight line with an origin ordinate very close to zero (0.017) and a slope very close to one (0.997). In other words, the descriptors predict the log P, log K^W, log K^S and I values of the different systems in an almost perfect way.

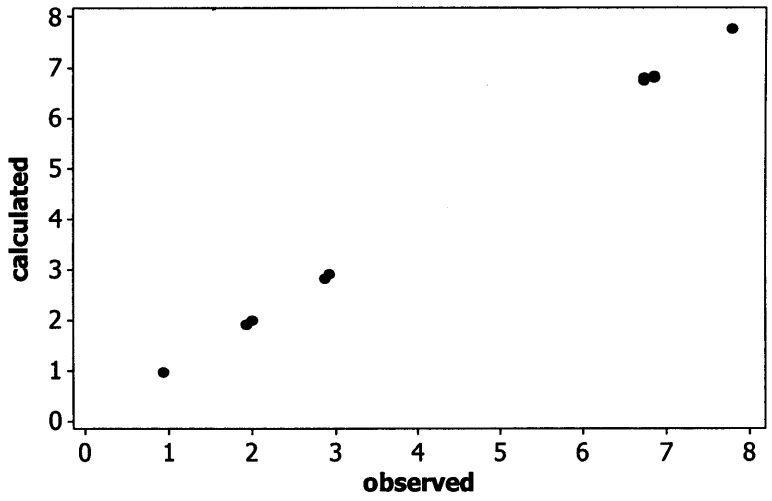


Figure 11.1. Plot of the calculated versus observed SP values for the systems in Table 11.2. Note that the two outliers have been removed from the graph.

The characteristic coefficients in Eqs. (11.1) and (11.2) for those systems used to determine the solvation properties of 1,2-dichlopropane are summarized in Tables 11.4 and 11.3, respectively.

Table 11.3
Characteristic coefficients in Eq. (11.2) for partitions between water and solvents

water/solvent system	*c*	*e*	*s*	*a*	*b*	*v*	Reference
gas phase	-0.994	0.577	2.549	3.813	4.841	-0.869	[21]
octanol	0.088	0.562	-1.054	0.034	-3.460	3.814	[22]
hexadecane	0.087	0.667	-1.617	-3.587	-4.869	4.433	[23]

Table 11.4
Characteristic coefficients in Eq. (11.1) for partitions between the gas phase and solvents

gas/solvent system	*c*	*e*	*s*	*a*	*b*	*l*	Reference
water	-1.271	0.822	2.743	3.904	4.814	-0.213	[21]
octanol	-0.198	0.002	0.709	3.519	1.429	0.858	[24]
hexadecane	0.000	0.000	0.000	0.000	0.000	1.000	[24]
SPB-1	0.836	-0.143	0.808	0.463	0.000	1.952	[17]
BP-1	0.819	0.000	0.671	1.007	0.000	1.960	[25]
BP-10	0.598	-0.445	2.637	3.217	-0.158	2.009	[20]
DB-1	0.668	0.000	0.597	1.358	0.208	1.995	[20]
Petrocol DH	0.579	-0.342	0.755	0.869	0.983	2.014	[26]

As mentioned above, it is important that the equations in Tables 11.3 and 11.4 contain different information that can be used to obtain descriptors. An approach of comparison of solvation equations is that of Ishihama and Asakawa [27]. The authors regard the five-descriptor equations as a line in five dimensions. Then for two equations, the angle, θ, between the lines is a measure of how close the equations are, mathematically. The larger the θ values, the less correlation there is between SP_i and SP_j. Identical equations would give a cos θ value of 1 (Table 11.5). In order to calculate values of θ, via Eq. (11.11) [27], for the 9 gas/solvent equations, one equation has to be taken as the standard and the first equation in Table 11.3 and Table 11.4 was chosen as the arbitrary standard.

$$\cos\theta_{ij} = \frac{e_i e_j + s_i s_j + a_i a_j + b_i b_j + l_i l_j}{\sqrt{e_i^2 + s_i^2 + a_i^2 + b_i^2 + l_i^2} + \sqrt{e_j^2 + s_j^2 + a_j^2 + b_j^2 + l_j^2}} \qquad (11.11)$$

The same applies for the three water/solvent equations, but in Eq. (11.11) instead of *l*, the coefficient system *v* must be used. Values of cos θ and θ are given in Table 11.5. As can be seen, there are huge variations in θ, and so the equations in Table 11.3 as well as those in Table 11.4 are sufficiently different to be able to determine descriptors for 1,2-dichloropropane.

Table 11.5
List of θ values

gas/solvent equations	cos θ	θ	water/solvent equations	cos θ	θ
water	1.000	0.000	gas phase	1.000	0.000
octanol	0.828	34.122	octanol	-0.625	128.673
BP-10	0.575	54.901	hexadecane	-0.863	149.625
Petrocol DH	0.546	56.896			
DB-1	0.441	63.831			
BP-1	0.340	70.105			
SPB-1	0.236	76.364			
hexadecane	-0.031	91.787			

In the same way as explained in this example, descriptors have been determined for a number of volatile organic compounds that have been found in indoor air, which is one

of the main purposes of this work. These VOCs and their descriptors are summarized in Table 11.6. On the other hand, in Table 11.7 descriptors for some other compounds of interest can be found. These descriptors have been obtained with the purpose of increasing the number of compounds for which the descriptors are known, in the SMART database (available at UCL).

Table 11.6
Determination of descriptors using log P, log K^W, log K and I values, when available, for some VOCs in table 11.12

Compund Name	E	S	A	B	L
1-hydroxy-2-propanone	0.359	0.77	0.29	0.73	2.432
1,2-dichloropropane	0.369	0.63	0.00	0.17	2.836
2,5-dimethylfuran	0.369	0.51	0.00	0.16	2.918
(E)-2-pentenal	0.383	0.75	0.00	0.43	3.045
γ-butyrolactone	0.387	1.38	0.00	0.65	3.331
1-octen-3-ol	0.260	0.39	0.35	0.62	4.168
2-methylclopentanone	0.494	0.82	0.00	0.56	3.485
2-ethylhexanoic acid	0.165	0.46	0.64	0.61	4.642
2,2,4,6,6-pentamethylheptane	0.000	0.00	0.00	0.00	4.592
sabinene	0.457	0.26	0.00	0.25	4.448
(E)-2-octenal	0.351	0.75	0.00	0.43	4.579
4-methyldecane	0.000	0.00	0.00	0.00	4.963
5-methyldecane	0.000	0.00	0.00	0.00	4.963
2-methyldecane	0.000	0.00	0.00	0.00	4.981
3-methyldecane	0.000	0.00	0.00	0.00	5.037
(E)-2-nonenal	0.351	0.75	0.00	0.43	5.050
(E)-2-undecenal	0.306	0.75	0.00	0.43	6.062
α-ylangene	0.281	0.24	0.00	0.33	6.351

Table 11.7
Determination of descriptors using log P, log K^W, log K and I values, when available, for some other VOCs

Compund Name	E	S	A	B	L
2-methylfuran	0.368	0.51	0.00	0.14	2.465
1-penten-3-ol	0.293	0.39	0.35	0.62	2.786

Table 11.7 *(continued)*

Compund Name	E	S	A	B	L
2-ethylfuran	0.361	0.49	0.00	0.15	2.921
1-hexen-3-ol	0.274	0.39	0.35	0.62	3.246
2-propylfuran	0.349	0.49	0.00	0.16	3.391
2-acetylfuran	0.629	1.11	0.00	0.54	3.685
1-hepten-3-ol	0.267	0.39	0.35	0.62	3.700
2-butylfuran	0.339	0.47	0.00	0.17	3.897
2-pentylfuran	0.333	0.47	0.00	0.17	4.384
2-hexylfuran	0.318	0.47	0.00	0.18	4.864
phenylacetic acid	0.730	1.08	0.66	0.57	4.962
2-heptylfuran	0.305	0.47	0.00	0.18	5.374
α-humulene	0.772	0.26	0.00	0.25	6.883
germacrene D	0.550	0.34	0.00	0.25	6.910
δ-cadinene	0.817	0.33	0.00	0.34	7.130

11.2.1 Statistical analysis

Because the analysis yields the one set of descriptors that best reproduce the observations, there are no statistics as regards standard deviation in the values of the descriptors. One can rectify this by adopting the method of 'leave-one-out', which is a particular case of cross-validation. Phenylacetic acid is taken as example, as the number of systems available for this compound is 31. The same procedure was used to obtain the descriptors of phenylacetic acid, as described previously. In this case, **E** was estimated on the base of values calculated from the refractive index. Log P values were available in twelve different water-solvent systems for this compound [12]. In this occasion log K^W was not available and was included as a variable to be calculated by Solver. This enables the corresponding nine log K^S values to be deduced through nine equations of type Eq. (11.10), once the log K^W is determined by Solver. In addition, Kovats retention indices, I, were available in eight GLC systems. These systems can be found in the references indicated in Table 11.8.

Knowing **E** = 0.730 and **V** = 1.0727, four descriptors, **S**, **A**, **B** and **L**, and log K^W, need to be calculated. With **S** = 1.08, **A** = 0.66, **B** = 0.57 and **L** = 4.962, 28 of the 31 initial data were reproduced with an SD of 0.084 log units.

The characteristic coefficients in Eqs. (11.1) and (11.2) for those systems used to determine the solvation properties of phenylacetic acid are summarized in Tables 11.10 and 11.9, respectively.

Table 11.8
Calculation of descriptors for phenylacetic acid

System	SP	SP_{calc}	SP_{obs}	Reference
water-octanol	log P	1.501	1.450	[12]
water-isobutanol	log P	1.532	1.430	[12]
water-pentanol	log P	1.625	1.570	[12]
water-dichloromethane	log P	0.411	0.600	[12]
water-trichloromethane	log P	0.449	0.420	[12]
water-heptane	log P	-1.427	-1.400	[12]
gas-water	log K^W	6.525	6.470	Solver
water-benzene	log P	-0.019	-0.030	[12]
water-toluene	log P	-0.112	-0.130	[12]
water-nitrobenzene	log P	0.740	0.670	[12]
water-diethyl ether	log P	1.186	1.330	[12]
water-dibutylether	log P	0.759	0.760	[12]
water-PGDP	log P	0.746	0.590	[12]
gas-octanol	log K^S	7.963	7.920	Eq. (11.10)
gas-dichloromethane	log K^S	6.982	7.070	Eq. (11.10)
gas-chloroform	log K^S	6.914	6.890	Eq. (11.10)
gas-heptane	log K^S	5.034	5.070	Eq. (11.10)
gas-water	log K^W	6.556	6.470	Solver
gas-benzene	log K^S	6.475	6.440	Eq. (11.10)
gas-toluene	log K^S	6.358	6.340	Eq. (11.10)
gas-nitrobenzene	log K^S	7.125	7.140	Eq. (11.10)
gas-diethyl ether	log K^S	7.647	7.800	Eq. (11.10)
gas-dibutylether	log K^S	7.265	7.230	Eq. (11.10)
GLC-DB-Wax	I/100	25.598	25.710	[28]
GLC-DB-Wax	I/100	25.151		[29]
GLC-DB-5MS	I/100	13.177		[30]
GLC-CP-SIL-5CB	I/100	12.218	12.120	[31]

Table 11.8 *(continued)*

System	SP	SP_{calc}	SP_{obs}	Reference
GLC-DB-Wax	I/100	25.399	25.340	[32]
GLC-HP-1	I/100	12.296	12.360	[33]
GLC-InnoWax	I/100	24.621		[33]
GLC- CP-SIL-5CB	I/100	12.390	12.300	[34]

Table 11.9
Characteristic coefficients in Eq. (11.2) for partitions between water and solvents

water/solvent system	*c*	*e*	*s*	*a*	*b*	*v*	Reference
gas phase	-0.994	0.577	2.549	3.813	4.841	-0.869	[21]
octanol	0.088	0.562	-1.054	0.034	-3.460	3.814	[22]
isobutanol	0.249	0.480	-0.639	-0.050	-2.284	2.758	[35]
pentanol	0.175	0.575	-0.787	0.020	-2.837	3.249	[36]
dichloromethane	0.314	0.001	0.022	-3.238	-4.137	4.259	[37]
trichloromethane	0.327	0.157	-0.391	-3.191	-3.437	4.191	[38]
heptane	0.325	0.670	-2.061	-3.317	-4.733	4.543	[39]
benzene	0.142	0.464	-0.588	-3.099	-4.625	4.491	[24]
toluene	0.143	0.527	-0.720	-3.010	-4.824	4.545	[24]
nitrobenzene	-0.181	0.576	0.003	-2.356	-4.420	4.263	[24]
diethyl ether	0.248	0.561	-1.016	-0.226	-4.553	4.075	[40]
dibutylether	0.252	0.677	-1.506	-0.807	-5.249	4.815	[40]
PGDP	0.139	0.376	-0.715	-1.034	-4.852	4.245	[35]

Table 11.10
Characteristic coefficients in Eq. (11.1) for partitions between the gas phase and solvents

gas/solvent system	*c*	*e*	*s*	*a*	*b*	*l*	Reference
octanol	-0.198	0.002	0.709	3.519	1.429	0.858	[24]
dichloromethane	0.121	-0.450	1.677	0.404	0.786	0.940	[24]
chloroform	0.116	-0.467	1.203	0.138	1.432	0.994	[38]
heptane	0.275	-0.162	0.000	0.000	0.000	0.983	[39]
water	-1.271	0.822	2.743	3.904	4.814	-0.213	[21]
benzene	0.107	-0.313	1.053	0.457	0.169	1.020	[21]
toluene	0.121	-0.222	0.938	0.467	0.099	1.012	[21]
nitrobenzene	-0.273	0.039	1.803	1.231	0.000	0.929	[21]

Table 11.10 *(continued)*

gas/solvent system	*c*	*e*	*s*	*a*	*b*	*l*	Reference
diethyl ether	0.206	-0.169	0.873	3.402	0.000	0.882	[40]
dibutylether	0.369	-0.216	0.026	2.626	-0.499	1.124	[41]
GLC-DB-Wax	-2.488	0.000	8.261	11.077	2.463	2.106	[28]
GLC-DB-Wax	-1,301	0,000	7,705	11,075	0,000	2,181	[29]
GLC-DB-5MS	-0,254	-0,571	1,228	1,947	2,407	1,988	[30]
GLC- CP-SIL-5CB	0.572	0.000	1.006	1.348	-0.654	2.024	[31]
GLC-DB-Wax	-1.307	0.523	9.055	10.206	0.000	1.977	[32]
GLC-HP-1	0.477	-0.574	1.064	1.465	0.000	2.040	[33]
GLC-Innowax	-0,707	0,931	5,690	10,871	1,090	2,158	[33]
GLC- CP-SIL-5CB	0.045	-0.609	1.664	1.596	-0.555	2.067	[34]

In the 'leave-one-out' method, the first equation is left out and a set of descriptors is calculated from equations 2–28. Then the second equation is left out and a set of descriptors calculated from equations 1, 3–28. Each equation is left out in turn, and so 28 sets of descriptors are obtained from which the statistics shown in Table 11.11 can be derived. The 'leave-one-out' values are almost exactly the same as those obtained from the entire 28 solvent set, but the internal self-consistency of the calculations can now be assessed, through, for example, the sd values for each individual descriptor. These SD values are very small indeed.

Table 11.11
Calculation of solvation descriptors by the method of 'leave-one-out'

Descriptor	S	A	B	L
Max. value[a]	1.09	0.67	0.57	4.969
Min. value[a]	1.07	0.65	0.57	4.938
Mean value[a]	1.08	0.66	0.57	4.961
SD[a]	**0.01**	**0.01**	**0.00**	**0.007**
All equations[b]	1.08	0.66	0.57	4.962

[a]By the 'leave-one-out' method. [b]All 28 equations used.

11.3. REFERENCES

1. M. J. Kamlet, R. M. Doherty, J.-L. M. Abboud, M. H. Abraham, R. W. Taft, CHEMTECH 16 (1986) 566
2. M. H. Abraham, R. M. Doherty, M. J. Kamlet, R. W. Taft, Chem. Br. 22 (1986) 551
3. M. H. Abraham, G. S. Whiting, R. M. Doherty, W. J. Shuely, J. Chem. Soc. Perkin Trans. 2, (1990) 1451
4. M. H. Abraham, G. S. Whiting, R. M. Doherty, W. J. Shuely, J. Chem. Soc. Perkin Trans. 2, 1990, 1851
5. M. H. Abraham, G. S. Whiting, R. M. Doherty, W. J. Shuely, J. Chromatogr. 587 (1991) 213
6. M. H. Abraham, J. Phys. Org. Chem. 6 (1993) 660
7. M. H. Abraham, Pure Appl. Chem. 65 (1993) 2503
8. M. H. Abraham, Chem. Soc. Rev. 22 (1993) 73
9. M. H. Abraham, A. M. Zissimos, W. E. Acree, Jr, Phys. Chem. Chem. Phys. 3 (2001) 3732
10. C. M. Du, K. Valko, C. Bevan, D. Reynolds, M. H. Abraham, J. Liq. Chromatogr. Relat. Technol. 24 (2001) 635
11. M. H. Abraham, J. C. McGowan, Chromatographia 23 (1987) 243
12. The MedChem Data Base 2005, BioByte. Corp., Pomona College, Daylight Chemical Information Systems, 27401 Los Altos, CA
13. M. H. Abraham, P. L. Grellier, R. A. McGill, J. Chem. Soc., Perkin Trans. 2 (1987) 797
14. M. H. Abraham, G.S. Whiting, R. M. Doherty, W. J. Shuely, J. Chromatogr. 587 (1991) 213
15. M. H. Abraham, G.S. Whiting, J. Chromatogr. 594 (1992) 229
16. M. H. Abraham, J. Chromatogr. 644 (1993) 95
17. R. J. Flanagan, P. J. Streete, J. D. Ramsey (1997) *Volatile Substance Abuse*, UNODC technical series, Number 5
18. G. Sun, P. Stremple, "Characterization of flavour, fragrance and many other compounds on DB-1 and DB-XLB", Available at: http://www.chem.agilent.com/ cag/cabu/pdf/ b-0279.pdf (Accessed: May 31, 2006)

19. Health and Safety Executive, (2000) *Volatile organic compounds in indoor air (4),* MDHS 96, HSE Books, ISBN 0 7176 1756 4
20. P. Ciccioli, A. Cecinato, E. Brancaleoni, M. Frattoni, A. Liberti, J. High. Resolut. Chromatogr. 15 (1992) 75
21. M. H. Abraham, J. Andonian-Haftvan, G. S. Whiting, A. Leo, R. W. Taft, J. Chem. Soc., Perkin Trans. 2 (1994) 1777
22. M. H. Abraham, J. Le, W. E. Acree, Jr., P. W. Carr, A. J. Dallas, Chemosphere 44 (2001) 855
23. M. H. Abraham, H. S. Chadha, G. S. Whiting, R. C. Mitchell, J. Pharm. Sci. 83 (1994) 1085
24. M. H. Abraham, unpublished equation
25. Health and Safety Executive, (1997) *Volatile organic compounds in air,* MDHS 88, HSE Books, ISBN 0 7176 2401 3
26. C. M. White, J. Hackett, R. R. Anderson, S. Kail, P. S. Spock, J. High. Resolut. Chromatogr. 15 (1992) 105
27. Y. Ishihama, N. Asakawa, J. Pharm. Sci. 88 (1999) 1305
28. M. Aznar, R. López, J. F. Cacho, V. Ferreira, J. Agric. Food Chem. 49 (2001) 2924
29. V. Ferreira, M. Aznar, R. López, J. Cacho, J. Agric. Food Chem. 49 (2001) 4818
30. M. E. Carunchia Whetstine, K. R. Cadwallader, M. Drake, J. Agric. Food Chem. 53 (2005) 3126
31. C. Guyot, A. Bouseta, V. Scheirman, S. Collin, J. Agric. Food Chem. 46 (1998) 625
32. Y. Kotseridis, R. Baumes, J. Agric. Food Chem. 48 (2000) 400
33. P. K. C. Ong, T. E. Acree, E. H. Lavin, J. Agric. Food Chem. 46 (1998) 611
34. C. Guyot, V. Scheirman, S. Collin, Food Chem. 64 (1999) 3
35. M. H. Abraham, J. Phys. Org. Chem. 6 (1994) 660
36. M. H. Abraham, H. S. Chadha, J. Dixon, A. J. Leo, J. Phys. Org. Chem. 7 (1994) 712
37. M. H .Abraham, K. Takacs-Novak, R. C. Mitchell, J. Pharm. Sci. 86 (1997) 310
38. M. H. Abraham, J.A. Platts, A. Hersey, A. Leo, R. W. Taft, J. Pharm. Sci. 88 (1999) 670
39. M. H. Abraham, C. E. Green, unpublished work
40. M. H. Abraham, A. M. Zissimos, W. E. Acree, Jr., New J. Chem. 27 (2003) 1041

41. M. H. Abraham, A. M. Zissimos, W. E. Acree, Jr., PCCP 3 (2001) 3732
42. Report No. 18: "Evaluation of VOC emissions from building products: solid flooring materials" (EUR 17334 EN).
43. Saarela *et al.* "Preliminary data base for material emissions", NKB Committee and Work Reports (1994).
44. P. Wolkoff, P.A. Nielsen, (1993) "Indoor Climate Labeling of Building Materials", Copenhagen, National Institute of Occupational Health, Danish Building Research Institute.
45. K. Sexton, J. L. Adgate, G. Ramachandran, G. C. Pratt, S. J. Mongin, T. H. Stock, M. T. Morandi, Environ. Sci. Technol. 38 (2004) 423.
46. C.-J. Lu, J. Whiting, R. D. Sacks, E. T. Zellers, Anal. Chem. 75 (2003) 1400.
47. M. Rehwagen, U. Schlink, O. Herbarth, Indoor Air 13 (2003) 283.
48. R. Meininghaus, E. Uhde, Indoor Air 12 (2002) 215.
49. S. K. Brown, Indoor Air 9 (1999) 209.
50. P. Wolkoff, G.D. Nielsen, Atmos. Environ. 35 (2001) 4407.
51. Inside IAQ, EPA/600N-98/002.
52. L. Mølhave, G. Clausen, B. Berglund, J. De Ceaurriz, A. Kettrup, T. Lindvall, M. Maroni, A. C. Pickering, U. Risse, H. Rothweiler, B. Seifert, M. Younes, Indoor Air 7 (1997) 225.
53. G. Bienie, S. Kurkiewicz, T. Wilczok, K. Klimek, L. Swiatkowska, A. Lusiak, J. Occup. Health 46 (2004) 181.
54. "Summary of Indoor and Outdoor Levels of Volatile Organic Compounds from Fule Oil Heated Homes in NYS, 1997-2003" New York State, Dept. of Health, Available at: http://www.health.state.ny.us/nysdoh/indoor/fuel_oil.htm (Accessed: October 14, 2004)
55. Health-related evaluation procedure for volatile organic compounds emissions from building products, Available at: www.umweltdaten.de/daten-e/agbb.pdf (Accessed: October 20, 2004)
56. D. Won and C.Y. Shaw (2004), "Investigation of building materials as VOCs sources in indoor air", Available at: http://www.ncree.org.tw/2004tcworkshop/pdf/20.pdf (Accessed: November 9, 2004)

11.4. APPENDIX

Table 11.12
List of VOCs identified in indoor air

Compound name	Reference
pentane	42
2-methylbutane	42
hexane	42
3-methylpentane	42
heptane	42
2-methylhexane	42
3-methylhexane	42
octane	42
nonane	42
2-methyloctane	42
3-methyloctane	42
decane	42
2-methylnonane	42
3,5-dimethyloctane	42
undecane	42
dodecane	42
tridecane	42
tetradecane	42
pentadecane	42
hexadecane	42
heptadecane	42
octadecane	42
eicosane	42
cyclohexane	42
methylcyclohexane	42
cis-1,4-dimethylcyclohexane	42
trans-1,4-dimethylcyclohexane	42
1,3-butadiene	42
dichloromethane	42
carbon tetrachloride	42
1,2-dichloroethane	42
vinyl chloride	42
trichloroethene	42
tetrachloroethene	42
dimethoxymethane	42
1,4-dioxane	42
formaldehyde	42
acetaldehyde	42

Table 11.12 *(continued)*

Compound name	Reference
propanal	42
butanal	42
pentanal	42
hexanal	42
heptanal	42
octanal	42
2-ethylhexanal	42
nonanal	42
decanal	42
trans-2-butenal	42
E-2-heptenal	42
Z-2-heptenal	42
acetone	42
2-butanone	42
3-methyl-2-butanone	42
4-methyl-2-pentanone	42
cyclopentanone	42
cyclohexanone	42
2-methylcyclohexanone	42
methyl formate	42
butyl formate	42
ethyl acetate	42
propyl acetate	42
isopropyl acetate	42
butyl acetate	42
isobutyl acetate	42
2-ethylhexyl acetate	42
vinyl acetate	42
2-methoxyethyl acetate	42
2-ethoxyethyl acetate	42
acetic acid	42
propanoic acid	42
isobutyric acid	42
pentanoic acid	42
2,2-dimethylpropanoic acid	42
hexanoic acid	42
heptanoic acid	42
octanoic acid	42
hexadecanoic acid	42
1-propanol	42
2-propanol	42
1-butanol	42

Table 11.12 *(continued)*

Compound name	Reference
2-methyl-1-propanol	42
2-methyl-2-propanol	42
1-pentanol	42
1-hexanol	42
1-octanol	42
2-ethyl-1-hexanol	42
cyclohexanol	42
2-methoxyethanol	42
2-ethoxyethanol	42
2-butoxyethanol	42
1,2-propanediol	42
benzene	42
toluene	42
ethylbenzene	42
o-xylene	42
m-xylene	42
p-xylene	42
propylbenzene	42
isopropylbenzene	42
1,2,3-trimethylbenzene	42
1,2,4-trimethylbenzene	42
1,3,5-trimethylbenzene	42
2-ethyltoluene	42
butylbenzene	42
1,2,4,5-tetramethylbenzene	42
1,3-diisopropylbenzene	42
1,4-diisopropylbenzene	42
styrene	42
α-methylstyrene	42
o-methylstyrene	42
m-methylstyrene	42
p-methylstyrene	42
ethynylbenzene	42
naphthalene	42
indene	42
1,4-dichlorobenzene	42
dimethyl phthalate	42
dibutyl phthalate	42
phenol	42
2,6-di-tert-butyl-4-methylphenol	42
furfural	42
pyridine	42

Table 11.12 *(continued)*

Compound name	Reference
3-ethylpyridine	42
1-methyl-2-pyrrolidinone	42
ε-caprolactam	42
α-pinene	42
β-pinene	42
limonene	42
camphene	42
(+-)-δ-3-carene	42
longifolene	42
α-cedrene	42
linalyl acetate	42
methyl methacrylate	42
1,4-dimethylcyclohexene	42
1-propenylbenzene	42
1-methyl-2-propylbenzene	42
2-phenyloctane	42
5-phenyldecane	42
5-phenylundecane	42
4-phenylcyclohexene	42
2,4,6-trimethyloctane	42
4-methyldecane	42
isododecane	42
2,2,4,6,6-pentamethylheptane	42
4,5-diethylnonane	42
phytane	42
pristine	42
TXIB	42
dimethoxyethane	42
(E)-2-nonenal	42
(Z)-2-decenal	42
2-undecenal	42
2-methylcyclopentanone	42
butyric acid	42
texanol	42
3-methyleneheptane	42
2-propenoic acid, 2-methyl-, heptyl ester	42
cis-1-methyl-4-(1-methylethyl)-cyclohexane	42
trans-1-methyl-4-(1-methylethyl)-cyclohexane	42
3-methylnonane	43
cis-1,2-dimethylcyclopentane	43
trans-1,2-dimethylcyclopentane	43
isopropylcyclohexane	43

Table 11.12 *(continued)*

Compound name	Reference
1-heptene	43
tridecene	43
1,3,5,7-cyclooctatetraene	43
2-heptanone	43
1-methyl-3-isopropylbenzene	43
1-methyl-4-isopropylbenzene	43
benzaldehyde	43
benzyl alcohol	43
β-myrcene	43
γ-terpinene	43
terpinolene	43
sabinene	43
caryophyllene	43
α-terpineol	43
2-(2-butoxyethoxy)-ethanol	43
2-propenoic acid, 6-methylheptyl ester	43
4-methyl-4-phenyl-2-pentanone	43
1,1-dimethylnonylbenzene	43
5-methylphenyl-2-hexanone	43
1,1-dimethyldecylbenzene	43
1,1-dimethyltetradecylbenzene	43
1,1,4,6,6-pentamethylheptylbenzene	43
pentadecylbenzene	43
1,1-dimethyldodecylbenzene	43
1,1-dimethylbutylbenzene	43
dimethylnonane	43
5-methyldecane	43
2-methyldecane	43
3-methyldecane	43
(-)-α-cubebene	43
longipinene	43
ylangene	43
junipene	43
2,6-dimethylnonane	43
methylpropylbenzene	43
ethyldimethylbenzene	43
methyl-x-(1-methylethenyl)-benzene	43
3,6-dimethyldecane	43
2-methylundecane	43
3-methylundecane	43
4-methylundecane	43
5-methylundecane	43

Table 11.12 *(continued)*

Compound name	Reference
(-)-β-pinene	43
methylcyclopentane	44
n-butyl propanoate	44
n-ethylacetamide	44
methyl isocyanate	44
ethanol (ethyl alcohol)	44
1,2-ethanediol (ethylene glycol)	44
1,4-dimethyl-x-ethylbenzene	44
hexamethylcyclotrisiloxane	44
ethoxyacetic acid	44
2-(2-butoxyethoxy)-ethanol acetate	44
bis-(1-methylpropyl) malonate	44
bis-(1-methylpropyl) succinate	44
bis-(1-methylpropyl) adipate	44
octamethyltrisiloxane	44
2-butanone oxime	44
octamethylcyclotetrasiloxane	44
tetraethyl orthosilicate	44
decamethyltetrasiloxane	44
decamethylcyclopentasiloxane	44
dodecamethylcyclohexasiloxane	44
tetradecamethylcycloheptasiloxane	44
hexamethylcyclotrisiloxane	44
chloroform (trichloromethane)	45
2-methylheptane	46
2,4-dimethylhexane	46
1,1,1-trichloroethane	46
2-hexanone	46
3-heptanone	46
3-octanone	46
3-methylbutanol	46
1-octen-3-ol	46
chlorobenzene	46
2-methylfuran	46
2,5-dimethylfuran	46
3-ethyltoluene	47
4-ethyltoluene	47
methyl benzoate	47
t-butyl acetate	48
chlorodifluoromethane	49
2-methylpentane	50
2-methylbut-1,3-diene	50

Table 11.12 *(continued)*

Compound name	Reference
chloromethane	50
dichlorofluoromethane	50
trichlorofluoromethane	50
hexachlorobenzene	50
2-phenoxyethanol	50
n-butoxy-2-propanol	50
hexyl acetate	51
2-methylbutan-1-ol	51
indan (indane)	51
2-methylbenzaldehyde	51
ethyl-3-ethoxypropionate	51
(2-methoxymethylethoxy)propanol	51
1-octene	52
1-decene	52
tetrahydrofuran	52
1-methoxypropan-2-ol	52
acetophenone	52
2-pentylfuran	52
3-methyl-3-butanone	52
acenaphthene	53
2-methylphenol	53
2,4-dimethylphenol	53
2,5-dimethylphenol	53
3,4-dimethylphenol	53
3,5-dimethylphenol	53
1-naphthol	53
2-naphthol	53
2,3-dimethylpentane	54
2,4-dimethylpentane	54
isooctane	54
ethylcyclohexane	54
cycloheptane	54
chloroethane	54
1,1-dichloroethane	54
1,1,2-trichloroethane	54
1,1,2,2-tetrachloroethane	54
1,2-dichloropropane	54
cis-1,2-dichloroethene	54
cis-1,3-dichloropropene	54
1,1-dichloroethene	54
trans-1,3-dichloropropene	54
hexachlorobuta-1,3-diene	54

Table 11.12 *(continued)*

Compound name	Reference
bromomethane	54
1,2-dibromoethane	54
difluorodichloromethane	54
1,1,2-trichlorotrifluoroethane	54
1,2-dichlorotetrafluoroethane	54
methyl tert-butyl ether	54
sec-butylbenzene	54
tert-butylbenzene	54
1,2-dichlorobenzene	54
1,3-dichlorobenzene	54
1,2,4-trichlorobenzene	54
ethyl methacrylate	54
1,2-dimethoxyethane	55
1,2-diethoxyethane	55
diglyme	55
ethylene carbonate	55
γ-butyrolactone	55
methyl acetate	55
dimethyl succinate	55
dimethyl glutarate	55
dimethyl adipate	55
2-ethylhexanoic acid	55
diethylene glycol	55
2-propoxyethanol	55
2-methoxypropanol	55
1,4-butanediol	55
4-hydroxy-4-methylpentan-2-one	55
1-hydroxy-2-propanone	55
1-isopropyl-2-methylbenzene	55
octylbenzene	55
decylbenzene	55
2,6-di-tert-butyl-4-methylphenol	55
1-isopropyl-4-methylcyclohexane	55
methyl acrylate	55
ethyl acrylate	55
n-butyl acrylate	55
2-ethylhexyl acrylate	55
2-(2-hexyloxyethoxy)ethanol	55
undecylbenzene	55
butyl glycollate	55
2-methylethoxyethanol	55
2-hexoxyethanol	55

Table 11.12 *(continued)*

Compound name	Reference
2-butoxyethyl acetate	55
2-methoxypropyl-1-acetate	55
propylene glycol diacetate	55
dipropylene glycol	55
dipropylene glycol monomethyl ether acetate	55
dipropylene glycol monopropyl ether	55
dipropylene glycol monobutyl ether	55
tripropyleneglycol monomethyl ether	55
triethylene glycol dimethyl ether	55
propylene glycol dimethyl ether	55
(E)-2-pentenal	55
2-octenal	55
glutaraldehyde	55
2-methoxy-1-methylethyl acetate	55
dibutyl fumarate	55
1,6-hexanediol diacrylate	55
dibutyl maleate	55
hexamethylenetetramine	55
tributyl phosphate	55
2-methyl-2h-isothiazol-3-one	55
propylcyclohexane	56
butylcyclohexane	56
dibutyl ether	56
α-terpinene	56
acrolein	56
decahydronaphthalene	56

Chapter 12 A New Algorithm for Nasal Pungency and Odour Detection Thresholds based on the Abraham Solvation Equation

12.0. INTRODUCTION

There is increasing concern over health impacts of indoor air pollution. Important pollutants are not only compounds such as formaldehyde, but volatile organic compounds (VOCs). VOCs are of crucial importance as regards air quality, especially indoor air quality. Most people spend more than 90% of their time indoors, where the concentration of VOCs in the atmosphere is typically from 2 to 20 times greater than concentrations found outdoors [1]. The perceived effect of VOCs can broadly be divided into odour and sensory irritation, the latter being so important that 40% of the workplace threshold limit values (TLVs) of the American Conference of Governmental Industrial Hygienists are based on this effect [2]. Sensory irritation includes both nasal pungency and eye irritation. Hence bioassays for nasal pungency are particularly relevant to the assessment of indoor air quality.

The VOCs that could be encountered at the work place or in the home number several thousand. In non-industrial buildings the number is less, although several hundred VOCs have been identified. These figures contrast with the number of VOCs actually tested by Cometto-Muñiz and Cain, used to obtained the algorithms for nasal pungency and odour detection threshold in this chapter. Nasal pungency thresholds of VOCs in man have been obtained for only 51 VOCs and odour detection threshold for only 64 by Cometto-Muñiz and Cain, using the systematic method described in Chapter 2 of this work.

The aim of this work is not only to apply, once again, the solvation equation of Abraham to the prediction of nasal pungency and odour detection threshold on humans, but to interpret the coefficients of each model regarding the physico-chemical properties of each system, and to demonstrate that the results of these equations lead to the conclusion that the main stage governing the process of the VOC to reach the receptor, is actually the transport of the VOC from the vapour phase into a solvent phase (mucus layer).

12.1. METHODOLOGY

A number of chemical features have been reported to correlate with sensory irritation. They include normal [3] and adjusted [4] boiling point, molecular weight [5], molecular geometry [5], saturated vapor pressure [5,6], Ostwald solubility coefficient [7], and other partition coefficients [8]. It has been pointed out that many of these quantitative structure–activity relationships (QSARs) provide limited chemical or mechanistic interpretations [9]. On the other hand, there have also been a number of early correlations of odour detection thresholds (ODTs) with various properties of odorants [10-13], but none have been very good statistically, and none have led to any conclusions of mechanistic significance. Two later studies related odour values to properties of homologous series of odorants. A study of ODT values [14] showed that for several homologous series, the ODT values could be correlated with the odorant activity coefficient in water [15]. Such correlations are of limited practical use, and of no mechanistic value.

In contrast, a recently described QSAR model, capable of describing and predicting the pungency and odour potency of airborne chemicals in humans, includes such interpretations. This model is based on the general solvation equation of Abraham [16]:

$$SP = c + e \cdot \mathbf{E} + s \cdot \mathbf{S} + a \cdot \mathbf{A} + b \cdot \mathbf{B} + l \cdot \mathbf{L} \qquad (12.1)$$

where SP is the dependent variable and represents some property (e.g., pungency or odour potency) of a series of chemical solutes (e.g., irritants or odorants) in a given (bio)phase solvent system (e.g., trigeminal nerve endings or odorant receptors). In this case, SP could be the reciprocal of the nasal pungency threshold as log (1/NPT) or odour detection threshold as log (1/ODT). The reciprocals are chosen simply because the larger the quantity, the more potent is the substance.

The five independent variables are: excess molar refraction (**E**), dipolarity/polarizability (**S**), overall or effective hydrogen-bond acidity (**A**), overall or effective hydrogen-bond basicity (**B**), and gas–liquid partition coefficient on hexadecane at 298 K (**L**). The term c and the coefficient for each independent variable (e, s, a, b, and l) are obtained by multiple-regression analysis. These coefficients have chemical and mechanistic meaning since they reflect the complementary properties that

the biophase must possess in order to be receptive to the irritant or odorant stimulus. In this way, the independent variables (i.e., **E**, **S**, **A**, **B**, and **L**) provide a physicochemical characterization of the stimulus (i.e., the irritant or odorant), whereas the corresponding coefficients (i.e., *e*, *s*, *a*, *b*, and *l*) provide a physicochemical characterization of the receptor area or biophase (e.g., trigeminal free nerve endings or the odour receptors) likely to interact with that stimulus [9]. The *e* coefficient measures the tendency of the biophase to interact with the irritant via polarizability-type interactions, mostly via π- and n-electron pairs. The *s* coefficient reflects the biophase dipolarity/polarizability (since a dipolar irritant will interact with a dipolar biophase, and a polarizable irritant will interact with a polarizable biophase). The *a* coefficient represents the complementary property to the irritant hydrogen-bond acidity, and thus is a measure of the biophase hydrogen-bond basicity (since an acidic irritant will interact with a basic biophase). Similarly, the *b* coefficient is a measure of the biophase hydrogen-bond acidity (since a basic irritant will interact with an acid biophase). Finally, the *l* coefficient is a measure of the biophase lipophilicity.

12.2. RESULTS

12.2.1 Algorithm for Nasal Pungency Thresholds

The nasal pungency thresholds used in this work are given in Table 12.1, as log (1/NPT) with NPT in ppm. The VOC descriptors are also included in Table 12.1.

Table 12.1
Descriptors, as defined in Eq. 12.1, and values of log(1/NPT) for volatile organic compounds (VOCs)

Stimuli	E	S	A	B	L	log (1/NPT)
dodecyl acetate	0.038	0.60	0.00	0.45	7.219	-0.097
octanoic acid	0.150	0.65	0.62	0.45	4.680	-0.290
decyl acetate	0.041	0.60	0.00	0.45	6.240	-0.700
hexanoic acid	0.174	0.63	0.62	0.44	3.697	-1.297
acetic acid	0.265	0.64	0.62	0.44	1.816	-1.623
menthol	0.400	0.50	0.23	0.58	5.177	-1.708
1-octanol	0.199	0.42	0.37	0.48	4.619	-1.750

Table 12.1 *(Continued)*

Stimuli	E	S	A	B	L	log (1/NPT)
butanoic acid	0.210	0.64	0.61	0.45	2.750	-1.790
octyl acetate	0.046	0.60	0.00	0.45	5.270	-1.919
1-heptanol	0.211	0.42	0.37	0.48	4.115	-2.320
2-nonanone	0.113	0.68	0.00	0.51	4.735	-2.341
1,8-cineole	0.383	0.33	0.00	0.76	4.688	-2.368
heptyl acetate	0.050	0.60	0.00	0.45	4.796	-2.490
formic acid	0.343	0.75	0.76	0.33	1.545	-2.500
4-heptanol	0.180	0.36	0.33	0.56	3.850	-2.523
linalool	0.398	0.55	0.20	0.67	4.794	-2.555
1-hexanol	0.210	0.42	0.37	0.48	3.610	-2.760
hexyl acetate	0.056	0.60	0.00	0.45	4.290	-2.863
2-heptanone	0.123	0.68	0.00	0.51	3.760	-2.952
p-cymene	0.607	0.49	0.00	0.19	4.590	-3.050
α-terpinene	0.526	0.25	0.00	0.15	4.715	-3.100
pyridine	0.631	0.84	0.00	0.52	3.022	-3.106
heptanal	0.140	0.65	0.00	0.45	3.865	-3.133
propyl benzene	0.604	0.50	0.00	0.15	4.230	-3.170
1-pentanol	0.219	0.42	0.37	0.48	3.106	-3.207
delta-3-carene	0.511	0.22	0.00	0.10	4.649	-3.213
pentyl acetate	0.067	0.60	0.00	0.45	3.844	-3.218
cumene	0.602	0.49	0.00	0.16	4.084	-3.218
octanal	0.160	0.65	0.00	0.45	4.361	-3.240
1-butanol	0.224	0.42	0.37	0.48	2.601	-3.476
2-pentanone	0.143	0.68	0.00	0.51	2.755	-3.493
butyl acetate	0.071	0.60	0.00	0.45	3.353	-3.544
sec-butyl acetate	0.044	0.57	0.00	0.47	3.054	-3.598
1-propanol	0.236	0.42	0.37	0.48	2.031	-3.693
hexanal	0.146	0.65	0.00	0.45	3.357	-3.700
2-butanol	0.217	0.36	0.33	0.56	2.338	-3.760
ethanol	0.246	0.42	0.37	0.48	1.485	-3.910
tert-butyl acetate	0.025	0.54	0.00	0.47	2.802	-3.978

Table 12.1 *(Continued)*

Stimuli	E	S	A	B	L	log (1/NPT)
ethyl benzene	0.613	0.51	0.00	0.15	3.778	-4.005
chlorobenzene	0.718	0.65	0.00	0.07	3.657	-4.023
propyl acetate	0.092	0.60	0.00	0.45	2.819	-4.245
2-propanol	0.212	0.36	0.33	0.56	1.764	-4.255
toluene	0.601	0.52	0.00	0.14	3.325	-4.473
1-octyne	0.155	0.22	0.09	0.10	3.521	-4.486
2-methyl-2-propanol	0.180	0.30	0.31	0.60	1.963	-4.515
methanol	0.278	0.44	0.43	0.47	0.970	-4.530
pentanal	0.163	0.65	0.00	0.45	2.851	-4.568
butanal	0.187	0.65	0.00	0.45	2.270	-4.775
ethyl acetate	0.106	0.62	0.00	0.45	2.314	-4.825
methyl acetate	0.142	0.64	0.00	0.45	1.911	-5.053
2-propanone	0.179	0.70	0.04	0.49	1.696	-5.118

The specific equation derived from Eq. 12.1 to describe and predict nasal pungency thresholds (NPTs) is

$$\log (1/\mathrm{NPT}) = - 8.217 + 1.690\,\mathbf{S} + 3.351\,\mathbf{A} + 1.095\,\mathbf{B} + 0.902\,\mathbf{L} \qquad (12.2)$$

with N = 49, R^2 = 0.951, SD = 0.273, and F = 213, where N is the number of compounds, R is the correlation coefficient, SD is the standard deviation, and F is the F-statistic. All other letters and symbols are as already defined for the general Eq. (12.1). In this NPT, the term $e \cdot \mathbf{E}$ from Eq. (12.2) did not achieve significance and was omitted, and two points were found to be outliers, i.e. acetic acid and linalool. The standard deviation in log(1/NPT) is only 0.27 log units in a range of 5 log units overall, and a plot of observed values of log(1/NPT) versus those predicted from Eq. (12.2) shows only random scatter about the line of identity, see Figure 12.1.

Observing the coefficients of Eq. (12.2), the trigeminal free nerve endings are revealed to not be able to interact with π- or n-electron pairs, as the term $e \cdot \mathbf{E}$ is not included in the equation. However, the rest of the coefficients will contribute positively to the predicted value of NPT. That means, that this system will be able to interact with any solute presenting a dipole, or a H-bond acid, or base, in its structure, or any

combination of them. And it will also be able to interact with solutes by dispersion forces, in other words, the energy required to create a given size cavity in this system is relatively low. These coefficients in Eq. (12.2) can be compared with those for various gas-condensed phase partitions that take place by simple transfer mechanisms, as shown in Table 12.2 [12-15]. There is considerable similarity between the NPT equation and equations for the solubility of gaseous VOCs in solvents such as wet 1-octanol and methanol. There is also some similarity with equations for the solubility of gaseous VOCs in a number of biophases [16].

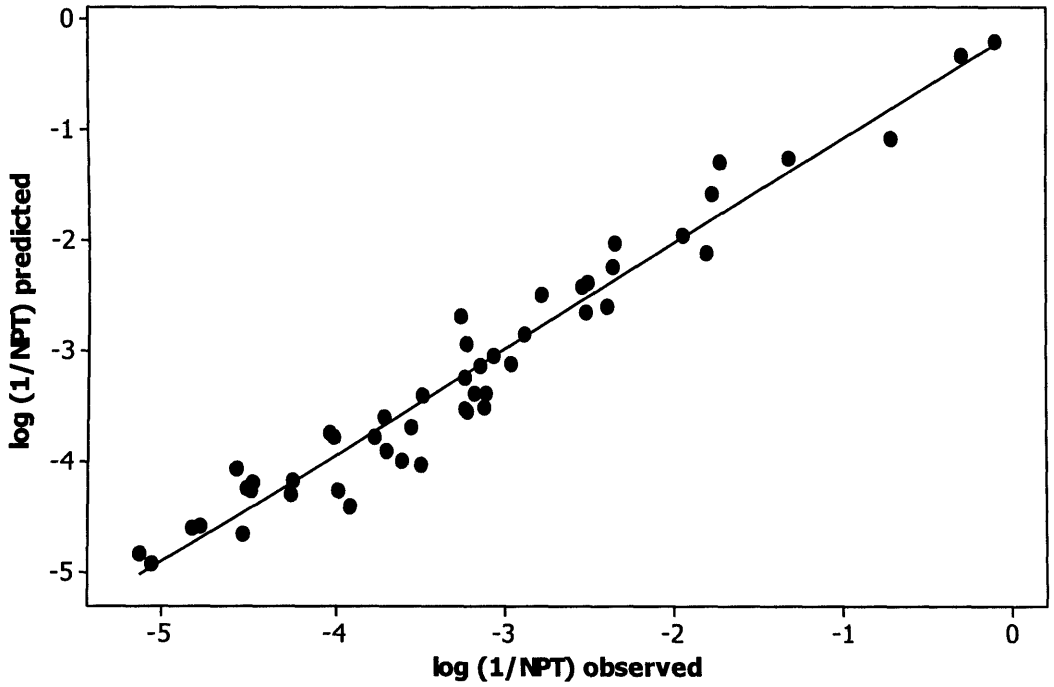


Figure 12.1. Plot of log(1/NPT) predicted vs log(1/NPT) observed from Eq. 12.2.

Table 12.2
Regression coefficients in Eq. (12.1) for gas-solvent phase partitions at 298 K

Phase	e	s	a	b	l
wet 1-octanol	0.002	0.709	3.519	1.429	0.858
dry methanol	-0.215	1.173	3.701	1.432	0.769
chloroform	-0.467	1.203	0.138	1.432	0.994
acetone	-0.277	1.522	3.258	0.078	0.863
dimethylformamide	-0.189	2.327	4.756	0.000	0.808
water	0.822	2.743	3.904	4.814	-0.213
brain[a]	0.427	0.286	2.781	2.787	0.609
muscle[a]	0.544	0.216	3.471	2.924	0.578

Table 12.2. *(Continued)*

Phase	e	s	a	b	l
fat[a]	-0.172	0.729	1.747	0.219	0.895
NPT, Eq. (12.2)[a]	0.000	1.69	3.35	1.09	0.902
ODT, Eq. (12.3)[a]	0.714	1.936	0.763	1.751	0.708

[a]At 310K

A good visual way of looking at the range of descriptors that this data set used to construct the NPT model covers is by plotting histograms of the values of the descriptors in the desired range. Five such histograms are in Figure 12.2, showing the frequency of the descriptors.

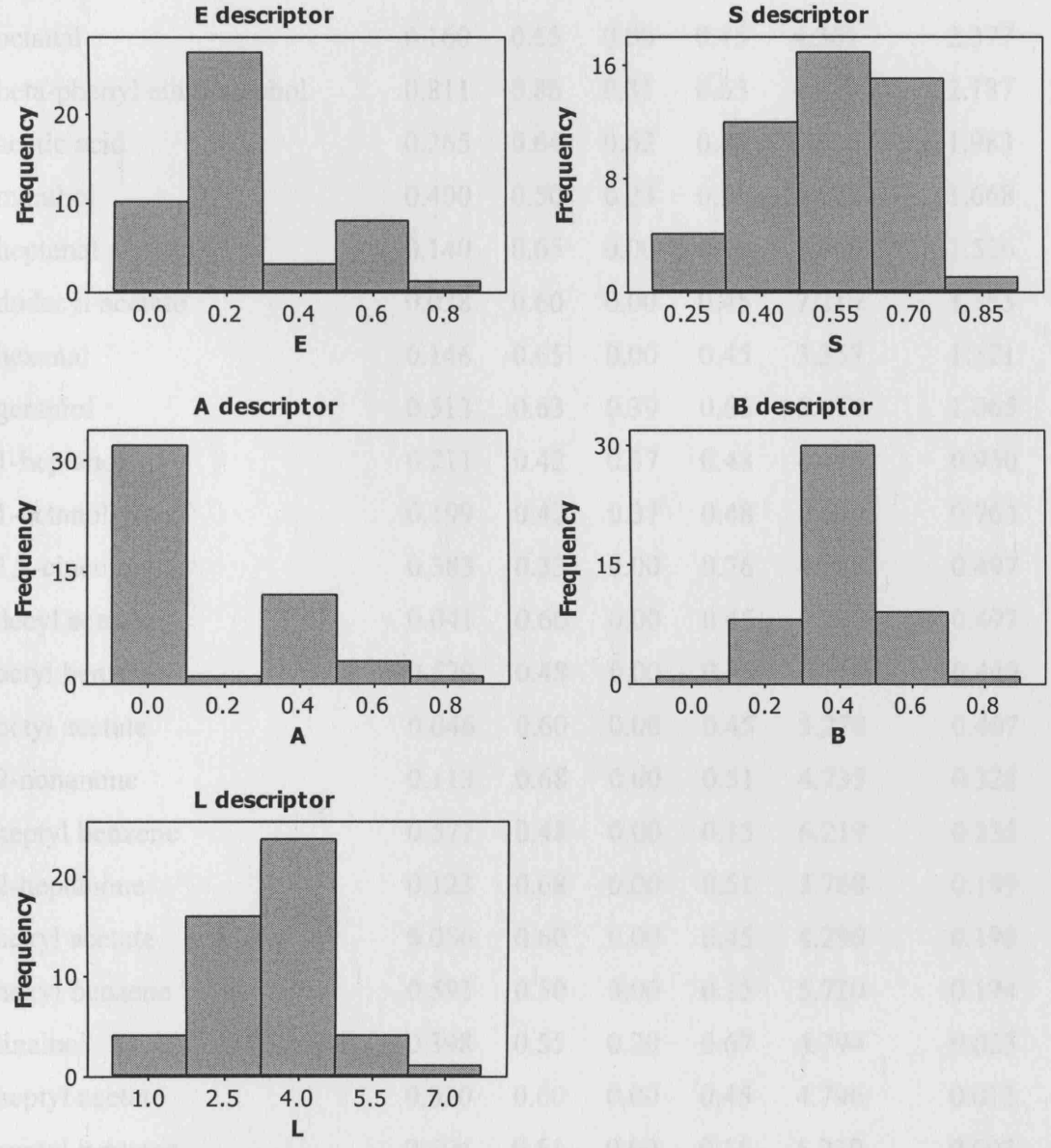


Figure 12.2. Histograms of the descriptors in the data set used to construct the nasal pungency threshold model.

12.2.2 Algorithm for Odour Detection Thresholds

The odour detection thresholds used in this work are given in Table 12.3, as log (1/ODT) with ODT in ppm. The VOC descriptors are also included in this table.

Table 12.3
Descriptors, as defined in Eq. 12.1, and values of log(1/ODT) for volatile organic compounds (VOCs)

Stimuli	E	S	A	B	L	log (1/ODT)
octanoic acid	0.150	0.65	0.62	0.45	4.680	4.941
hexanoic acid	0.174	0.63	0.62	0.44	3.697	2.585
butanoic acid	0.210	0.64	0.61	0.45	2.750	2.444
octanal	0.160	0.65	0.00	0.45	4.361	2.377
beta-phenyl ethyl alcohol	0.811	0.86	0.31	0.65	4.628	2.187
acetic acid	0.265	0.64	0.62	0.44	1.816	1.983
menthol	0.400	0.50	0.23	0.58	5.177	1.668
heptanal	0.140	0.65	0.00	0.45	3.865	1.526
dodecyl acetate	0.038	0.60	0.00	0.45	7.219	1.353
hexanal	0.146	0.65	0.00	0.45	3.357	1.121
geraniol	0.513	0.63	0.39	0.66	5.479	1.065
1-heptanol	0.211	0.42	0.37	0.48	4.115	0.950
1-octanol	0.199	0.42	0.37	0.48	4.619	0.764
1,8-cineole	0.383	0.33	0.00	0.76	4.688	0.497
decyl acetate	0.041	0.60	0.00	0.45	6.240	0.497
octyl benzene	0.579	0.48	0.00	0.15	6.714	0.440
octyl acetate	0.046	0.60	0.00	0.45	5.270	0.407
2-nonanone	0.113	0.68	0.00	0.51	4.735	0.328
heptyl benzene	0.577	0.48	0.00	0.15	6.219	0.255
2-heptanone	0.123	0.68	0.00	0.51	3.760	0.199
hexyl acetate	0.056	0.60	0.00	0.45	4.290	0.198
hexyl benzene	0.591	0.50	0.00	0.15	5.720	0.194
linalool	0.398	0.55	0.20	0.67	4.794	0.023
heptyl acetate	0.050	0.60	0.00	0.45	4.796	0.013
pentyl benzene	0.594	0.51	0.00	0.15	5.230	0.003

Table 12.3 *(Continued)*

Stimuli	E	S	A	B	L	log (1/ODT)
cumene	0.602	0.49	0.00	0.16	4.084	-0.033
pentyl acetate	0.067	0.60	0.00	0.45	3.844	-0.074
pyridine	0.631	0.84	0.00	0.52	3.022	-0.108
p-cymene	0.607	0.49	0.00	0.19	4.590	-0.118
α-terpinene	0.526	0.25	0.00	0.15	4.715	-0.150
propyl benzene	0.604	0.50	0.00	0.15	4.230	-0.154
1-pentanol	0.219	0.42	0.37	0.48	3.106	-0.160
delta-3-carene	0.511	0.22	0.00	0.10	4.649	-0.223
butyl acetate	0.071	0.60	0.00	0.45	3.353	-0.380
butanal	0.187	0.65	0.00	0.45	2.270	-0.474
butyl benzene	0.600	0.51	0.00	0.15	4.730	-0.628
(s)- (-)-limonene	0.488	0.28	0.00	0.21	4.725	-0.655
1-hexanol	0.210	0.42	0.37	0.48	3.610	-0.726
pentanal	0.163	0.65	0.00	0.45	2.851	-0.730
tert-butyl acetate	0.025	0.54	0.00	0.47	2.802	-0.755
formic acid	0.343	0.75	0.76	0.33	1.545	-0.883
4-heptanol	0.180	0.36	0.33	0.56	3.850	-0.915
ethyl benzene	0.613	0.51	0.00	0.15	3.778	-0.965
γ-terpinene	0.497	0.32	0.00	0.20	4.815	-0.990
(r)-(+)- limonene	0.488	0.28	0.00	0.21	4.725	-0.993
1-butanol	0.224	0.42	0.37	0.48	2.601	-1.012
beta-pinene	0.530	0.24	0.00	0.19	4.394	-1.070
chlorobenzene	0.718	0.65	0.00	0.07	3.657	-1.110
1-propanol	0.236	0.42	0.37	0.48	2.031	-1.250
α-pinene	0.446	0.14	0.00	0.12	4.308	-1.278
sec-butyl acetate	0.044	0.57	0.00	0.47	3.054	-1.330
2-pentanone	0.143	0.68	0.00	0.51	2.755	-1.348
propyl acetate	0.092	0.60	0.00	0.45	2.819	-1.398
ethanol	0.246	0.42	0.37	0.48	1.485	-1.728
2-butanol	0.217	0.36	0.33	0.56	2.338	-1.978
toluene	0.601	0.52	0.00	0.14	3.325	-2.102

Table 12.3 *(Continued)*

Stimuli	E	S	A	B	L	log (1/ODT)
1-octyne	0.155	0.22	0.09	0.10	3.521	-2.133
ethyl acetate	0.106	0.62	0.00	0.45	2.314	-2.238
1-octene	0.094	0.08	0.00	0.07	3.568	-2.313
2-propanol	0.212	0.36	0.33	0.56	1.764	-2.700
2-methyl-2-propanol	0.180	0.30	0.31	0.60	1.963	-2.783
methanol	0.278	0.44	0.43	0.47	0.970	-3.208
methyl acetate	0.142	0.64	0.00	0.45	1.911	-3.458
2-propanone	0.179	0.70	0.04	0.49	1.696	-4.078

The ranges of the descriptors in Table 12.3, including the frequency of them, are shown in Figure 12.3.

As a first step Eq. (12.1) was applied to all these VOCs except the carboxylic acids and aliphatic aldehydes that were clearly out of line, as shown in the initial analysis, section 12.2.3. The VOCs, 2-propanone, methyl acetate and 1-heptanol were then also revealed to be outliers, and were removed to yield to the correlation equation,

$$\log(1/\mathrm{ODT}) = -5.295 + 0.714\,\mathbf{E} + 1.936\,\mathbf{S} + 0.763\,\mathbf{A} + 1.751\,\mathbf{B} + 0.708\,\mathbf{L} \qquad (12.3)$$

with N = 51, R^2 = 0.813, SD = 0.523, and F = 39, where all letters and symbols are as already defined.

The coefficients in Eq. (12.3) are also very similar to those for transfer from the gas phase to organic solvents (Table 12.2). It therefore appears that simple transfer from the gas phase to a biophase must play a major role in the relationship of ODTs to the structure of VOCs, of the order of 81% of the total effect. This contrast with the corresponding equations for log (1/NPT) that account for 95% of the total effect.

In order to study what others factors, as well as simple transport, influence the ODT values, it is instructive to plot the residuals in Eq. (12.3) against the 'size' parameter, **L**. The residuals are not random, and both small VOCs and large VOCs are less potent than expected, see Figure 12.4. It can bee seen from this figure that the residuals follow a 'parabolic-term' curve. As molecular size increases, the residual values increases to a maximum value and then decreases. A similar analysis was carried out by Abraham and co-workers [17]. In that analysis, the same result was obtained, and

Abraham and co-workers suggested that the pattern of residuals in Figure 12.4 is due to an extra effect, in addition to simple transfer. The effect can simply be quantified and incorporated into Eq. (12.3) by addition of a parabolic term in $\mathbf{L}^2$,

$$\log(1/\mathrm{ODT}) = -6.981 + 0.393\,\mathbf{E} + 2.225\,\mathbf{S} + 1.572\,\mathbf{A} + 1.427\,\mathbf{B} + 1.604\,\mathbf{L} - 0.106\,\mathbf{L}^2 \qquad (12.4)$$

with N = 51, R^2 = 0.855, SD = 0.465, and F = 43. These statistics are actually quite good, bearing in mind the experimental error in the ODT values.

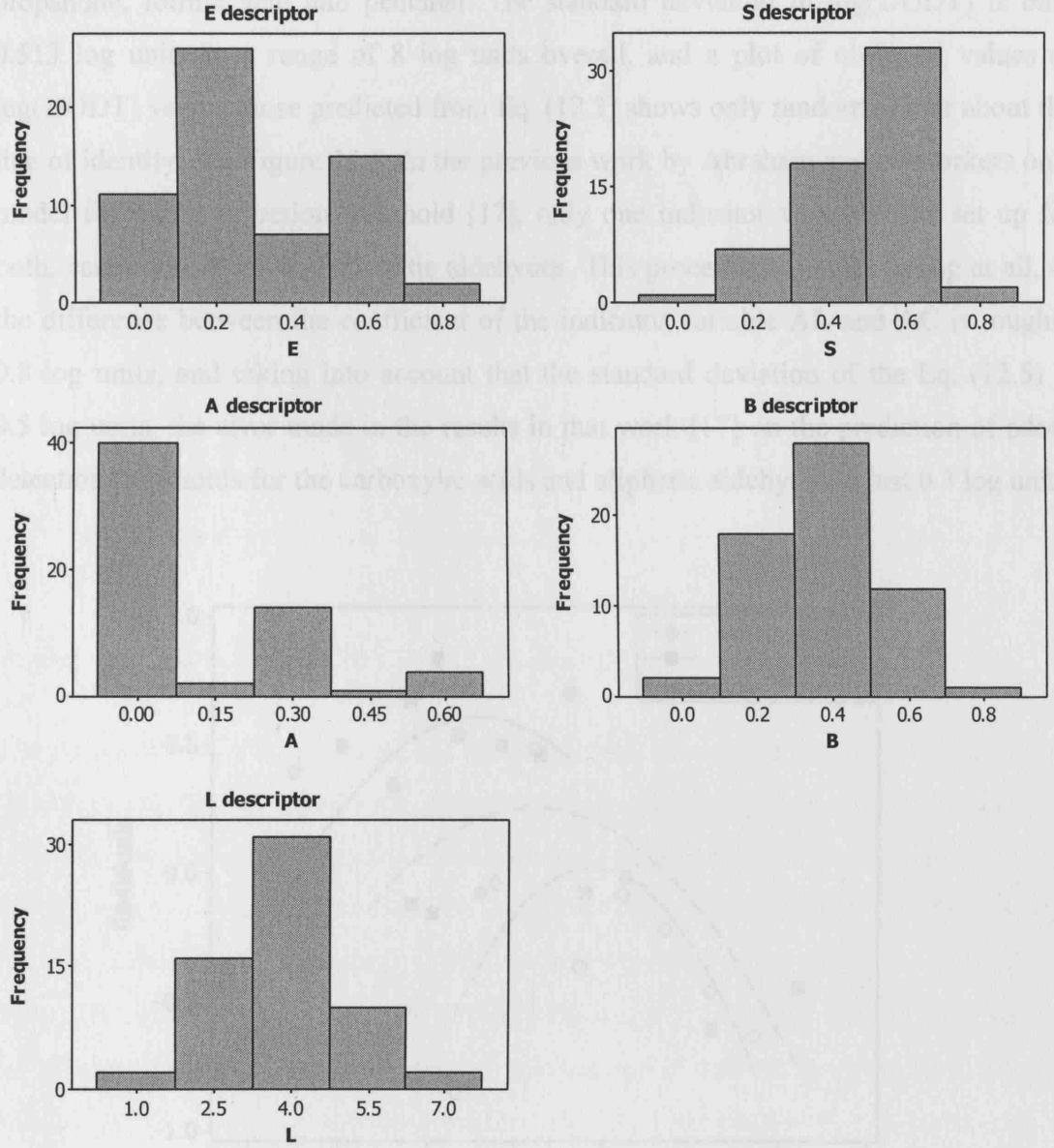


Figure 12.3. Histograms of the descriptors in the data set used to construct the odour detection threshold model.

Carboxylic acids and aliphatic aldehydes can now be included into Eq. (12.1) by means of an indicator variable for each of these groups. The final specific equation derived to describe and predict odour detection threshold (ODT) from this work is

$$\log(1/\mathrm{ODT}) = -7.475 + 0.283\,\mathbf{E} + 2.220\,\mathbf{S} + 2.226\,\mathbf{A} + 1.238\,\mathbf{B} + 1.821\,\mathbf{L} - 0.126\,\mathbf{L}^2 + 1.848\,\mathbf{AL} + 2.604\,\mathbf{AC} \quad (12.5)$$

with N = 61, R^2 = 0.905, SD = 0.508, and F = 62, where all letters and symbols are as already defined, except **AL** and **AC**. In this analysis, three outliers were found, i.e. 2-propanone, formic acid and pentanal. The standard deviation in log(1/ODT) is only 0.513 log units in a range of 8 log units overall, and a plot of observed values of log(1/ODT) versus those predicted from Eq. (12.3) shows only random scatter about the line of identity, see Figure 12.5. In the previous work by Abraham and co-workers on a model for odour detection threshold [17], only one indicator variable was set up for both, carboxylic acids and aliphatic aldehydes. This procedure was not wrong at all, as the difference between the coefficient of the indicator variable **AL** and **AC** is roughly 0.8 log units, and taking into account that the standard deviation of the Eq. (12.5) is 0.5 log units, the error made in the results in that work [17] on the prediction of odour detection thresholds for the carboxylic acids and aliphatic aldehydes is just 0.3 log units.

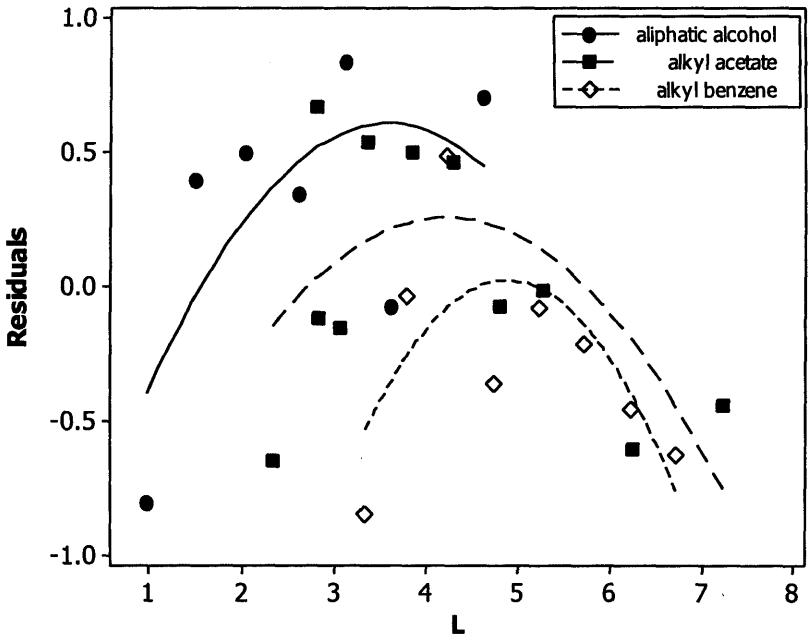


Figure 12.4. Residuals (observed-calculated values in Eq. (12.3)) against the L descriptor.

Observing the coefficients of Eq. (12.5), one can see that they will all contribute positively to the predicted value of ODT. That means, that this system will be able to interact with any solute presenting π- or n-electron pairs, a dipole, or a hydrogen-bond acid, or base, in its structure, or any combination of them. And it will also be able to interact with solutes by dispersion forces, in other words, the energy required to create a given size cavity in this system is also relatively low.

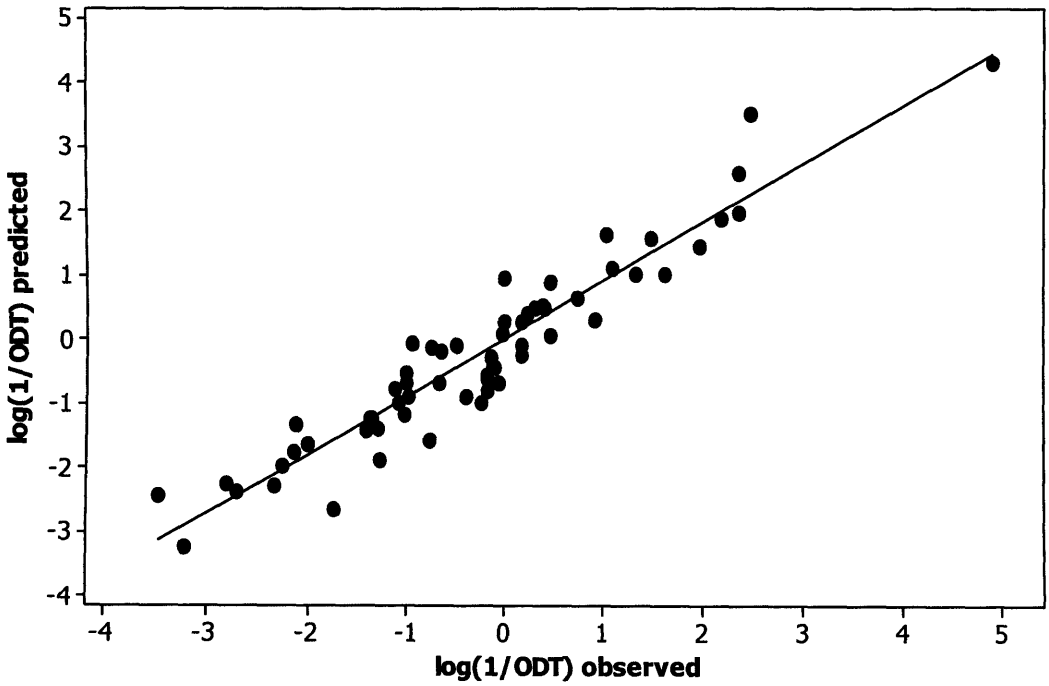


Figure 12.5. Plot of log(1/ODT) predicted vs log(1/ODT) observed from Eq. 12.3.

The indicator variables **AL** and **AC** are introduced to account for the larger values of aldehydes and acids, respectively, than predicted using Eq. (12.1) without the indicator variables. **AL** has the value 1 for aldehydes and zero for all other VOCs. Similarly, **AC** has the value 1 for acids and zero for all others VOCs. The indicator variables used in Eq. (12.5) are not just arbitrary variables used to obtain a better fit to the data; similar excess potency has been described before. Alarie et al. [18,19] investigated the sensory irritation of mice by airborne chemicals, and classed chemicals as acting by a physical mechanism (*p*) or by a chemical mechanism (*c*). Compounds that induced sensory irritation by a chemical mechanism were identified through an increase in potency by comparison with that calculated for irritation by a physical mechanism. In essence, this is the same procedure that we have used to identify compounds that are more potent

than calculated from Eq. (12.5). Alarie et al. [18,19] showed that carboxylic acids and aldehydes were more potent than expected, exactly as we have done.

12.2.3 Initial analysis

An analysis of the data in Table 12.3 as follows, will allow identification of those series of compounds that are entirely out of line, i.e. carboxylic acids and aliphatic aldehydes. The application of Eq (12.1) to all the VOCs, except 2-propanone and formic acid, revealed to be outliers, leads to Eq. (12.6)

$$\log (1/ODT) = - 5.558 + 0.522\,\mathbf{E} + 1.796\,\mathbf{S} + 1.507\,\mathbf{A} + 1.505\,\mathbf{B} + 0.807\,\mathbf{L} + 1.699\,\mathbf{AL} + 3.081\,\mathbf{AC} \quad (12.6)$$

with N = 62, $R^2 = 0.861$, SD = 0.605, F = 48.

For any homologous series, the descriptors **E**, **S**, **A** and **B** in Eq. (12.6) are almost constant except for **L**. Then Eq. (12.6) reverts to the simple equation,

$$\log(1/ODT) = c' + l\,\mathbf{L} \quad (12.7)$$

where c' is constant for any homologous series and is given by $c' = c + e\,\mathbf{E} + s\,\mathbf{S} + a\,\mathbf{A} + b\,\mathbf{B}$. Then for each homologous series a plot of log (1/ODT) against **L** will have a known slope of 0.807 and a given intercept of c'.

The graphs in Figure 12.6a, 12.6b, and 12.6c show that these data are scattered about the line calculated from Eq. (12.7), but nothing else. However, in Figure 12.6d one can also see that one ketone, 2-propanone, is out of line by some 2 log units. The graph shown in Figure 12.6e for the aliphatic aldehydes is quite different. Now all the aldehydes are out of line, and are all more potent than calculated, in an average of 1.7 log units. The best line through the observed data points is parallel to the calculated one from Eq. (12.7), suggesting that the indicator variable for aldehydes will bring them all into line. A similar situation is found for the aliphatic carboxylic acids, Figure 12.6f, were they are all in average 3.1 log units more potent than calculated. Moreover, in this graph one acid, formic acid, is out of line. Note that in Figure 12.6, the calculated lines all have the same slope. This may not be apparent from the various figures as they are plotted on different scales.

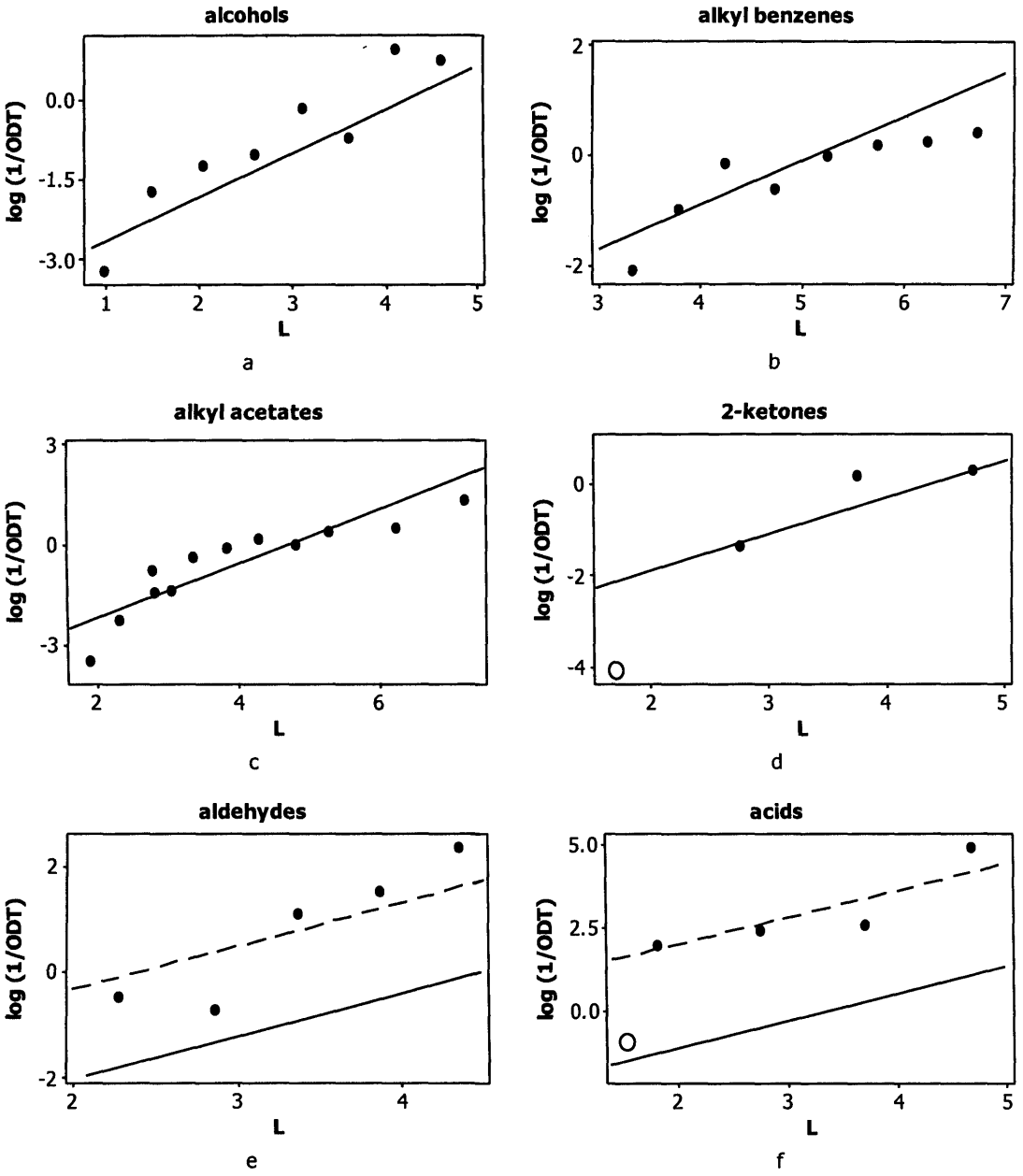


Figure 12.6. Scatterplot of log (1/ODT) against **L** for eight different homologous series of VOCs. Calculated line without the indicator variable (———). Calculated line with the indicator variable (– – – –). Outliers (o).

12.3. ESTIMATION OF NASAL PUNGENCY AND ODOUR DETECTION THRESHOLDS

All the descriptors in Eqs. (12.2) and (12.5) for a variety of VOCs are known, so that it is then possible to calculate log (1/NPT) and log (1/ODT) values. The corresponding log (1/NPT) and log (1/ODT) values are in Table 12.4 for VOCs that have been found in indoor air, see the appendix in Chapter 11. The compounds are ordered by the

log (1/NPT), in ascending order. As mentioned, the larger the quantity, the more potent is the VOC. Therefore, in Table 12.4 the order of pungency (or irritancy) of the VOC irritants is then shown.

Table 12.4
Calculated values of nasal pungency and odour detection thresholds for VOCs

VOC	log (1/NPT)	log (1/ODT)
difluorodichloromethane	-7.000	-5.311
1,2-dichlorotetrafluoroethane	-6.845	-4.968
chlorodifluoromethane	-6.502	-5.278
2-methylbutane	-6.401	-4.320
methyl chloride	-6.354	-4.403
1,3-butadiene	-6.327	-4.240
pentane	-6.267	-4.127
vinyl chloride	-6.254	-4.188
formaldehyde	-6.149	-2.518
trichlorofluoromethane	-5.976	-3.725
2-methylpentane	-5.959	-3.706
chloroethane	-5.918	-3.698
bromomethane	-5.911	-3.650
3-methylpentane	-5.889	-3.614
2-methylbut-1,3-diene	-5.824	-3.482
hexane	-5.810	-3.513
2,4-dimethylpentane	-5.683	-3.354
2-methylhexane	-5.510	-3.145
2,3-dimethylpentane	-5.497	-3.129
methyl formate	-5.493	-3.309
acetaldehyde	-5.482	-1.474
3-methylhexane	-5.471	-3.099
methylcyclopentane	-5.426	-2.960
isooctane	-5.415	-3.035
cyclohexane	-5.374	-2.876
heptane	-5.355	-2.966
trans-1,2-dimethylcyclopentane	-5.253	-2.765

Table 12.4 *(Continued)*

VOC	log (1/NPT)	log (1/ODT)
1-heptene	-5.242	-2.789
2,4-dimethylhexane	-5.223	-2.819
dimethoxymethane	-5.162	-2.785
methylcyclohexane	-5.122	-2.617
cis-1,2-dimethylcyclopentane	-5.096	-2.573
2-methylheptane	-5.078	-2.664
methyl tert-butyl ether	-5.069	-2.651
dichloromethane	-5.043	-2.653
carbon tetrachloride	-5.028	-2.365
2-methylfuran	-5.027	-2.405
propanal	-4.989	-0.681
1,1,1-trichloroethane	-4.960	-2.313
methyl acetate	-4.919	-2.437
octane	-4.900	-2.483
1,1-dichloroethane	-4.855	-2.408
acetone	-4.834	-2.449
1-octene	-4.787	-2.291
vinyl acetate	-4.723	-2.123
cycloheptane	-4.707	-2.138
benzene	-4.672	-1.879
2-methyloctane	-4.640	-2.235
chloroform	-4.627	-2.167
3-methyloctane	-4.611	-2.209
trans-1,4-dimethylcyclohexane	-4.597	-2.019
ethyl acetate	-4.589	-1.972
trichloroethene	-4.587	-1.964
butanal	-4.578	-0.090
cis-1,2-dichloroethene	-4.563	-1.999
ethylcyclohexane	-4.551	-2.012
2,5-dimethylfuran	-4.548	-1.799
methyl acrylate	-4.513	-1.822

Table 12.4 *(Continued)*

VOC	log (1/NPT)	log (1/ODT)
cis-1,4-dimethylcyclohexane	-4.509	-1.929
nonane	-4.445	-2.063
isopropyl acetate	-4.443	-1.793
tetrahydrofuran	-4.435	-1.720
2-butanone	-4.413	-1.737
1,2-dichloropropane	-4.408	-1.611
ethanol	-4.402	-2.629
3,5-dimethyloctane	-4.375	-2.005
1,2-dichloroethane	-4.359	-1.726
2-propanol	-4.298	-2.368
methyl methacrylate	-4.276	-1.529
t-butyl acetate	-4.262	-1.574
2-methyl-2-propanol	-4.244	-2.236
tetrachloroethene	-4.241	-1.409
isopropylcyclohexane	-4.227	-1.745
2-methylnonane	-4.200	-1.865
ethyl acrylate	-4.188	-1.410
toluene	-4.186	-1.315
3-methylnonane	-4.171	-1.842
propyl acetate	-4.168	-1.428
propylcyclohexane	-4.158	-1.670
3-methyl-2-butanone	-4.132	-1.374
butyl formate	-4.068	-1.288
pentanal	-4.054	0.587
trans-2-butenal	-3.999	0.725
decane	-3.990	-1.709
α-pinene	-3.963	-1.383
ethyl methacrylate	-3.960	-1.181
1,2-dimethoxyethane	-3.946	-1.168
1-isopropyl-4-methylcyclohexane	-3.917	-1.509
1-decene	-3.916	-1.519

Table 12.4 *(Continued)*

VOC	log (1/NPT)	log (1/ODT)
1-propanol	-3.910	-1.879
isobutyl acetate	-3.888	-1.116
1,1-dichloroethene	-3.884	-0.899
trans-1,3-dichloropropene	-3.884	-0.899
cis-1,3-dichloropropene	-3.869	-0.876
ethylbenzene	-3.783	-0.902
camphene	-3.779	-1.161
4-methyl-2-pentanone	-3.774	-0.946
n-butyl ether	-3.762	-1.157
chlorobenzene	-3.743	-0.768
butylcyclohexane	-3.713	-1.339
1,3,5,7-cyclooctatetraene	-3.703	-0.773
m-xylene	-3.700	-0.812
p-xylene	-3.700	-0.815
butyl acetate	-3.686	-0.877
1,4-dioxane	-3.640	-0.712
β-pinene	-3.640	-0.988
2-methyl-1-propanol	-3.616	-1.469
hexanal	-3.598	1.108
2-hexanone	-3.545	-0.672
o-xylene	-3.542	-0.628
3-carene	-3.542	-0.976
undecane	-3.535	-1.417
isopropylbenzene	-3.530	-0.683
1,1,2-trichloroethane	-3.522	-0.747
pyridine	-3.502	-0.435
sabinene	-3.492	-0.852
styrene	-3.465	-0.445
β-myrcene	-3.430	-0.785
1-butanol	-3.396	-1.177
propylbenzene	-3.392	-0.560

Table 12.4 *(Continued)*

VOC	log (1/NPT)	log (1/ODT)
α-terpinene	-3.377	-0.801
1,2-dibromoethane	-3.361	-0.426
3-ethyltoluene	-3.302	-0.460
n-butyl propanoate	-3.299	-0.505
n-butyl acrylate	-3.291	-0.437
4-ethyltoluene	-3.289	-0.449
cyclopentanone	-3.289	-0.258
2-pentylfuran	-3.282	-0.565
limonene	-3.252	-0.664
ethynylbenzene	-3.242	-0.425
1,3,5-trimethylbenzene	-3.212	-0.369
tert-butylbenzene	-3.211	-0.407
2-methoxyethanol	-3.201	-0.828
2-ethyltoluene	-3.170	-0.325
sec-butylbenzene	-3.166	-0.394
heptanal	-3.140	1.569
3-heptanone	-3.137	-0.269
1,2-diethoxyethane	-3.133	-0.227
2-heptanone	-3.118	-0.234
γ-terpinene	-3.114	-0.529
dodecane	-3.079	-1.191
2-methylbutan-1-ol	-3.077	-0.789
3-methylbutanol	-3.077	-0.796
m-cymene	-3.071	-0.295
1,2,4-trimethylbenzene	-3.057	-0.203
α-methylstyrene (2-phenylpropene)	-3.056	-0.083
p-cymene	-3.041	-0.276
2-methoxyethyl acetate	-3.027	-0.044
o-methylstyrene	-2.996	-0.012
1,3-dichlorobenzene	-2.984	-0.010
m-methylstyrene	-2.975	-0.009

Table 12.4 *(Continued)*

VOC	log (1/NPT)	log (1/ODT)
terpinolene	-2.966	-0.418
p-methylstyrene	-2.954	0.010
1-hydroxy-2-propanone	-2.951	-0.431
o-cymene	-2.944	-0.149
1-pentanol	-2.940	-0.622
acetic acid	-2.938	1.441
1,4-dichlorobenzene	-2.927	0.046
butylbenzene	-2.924	-0.193
2-ethylhexanal	-2.907	1.758
2-ethoxyethanol	-2.894	-0.459
indan	-2.865	0.026
1,2,3-trimethylbenzene	-2.860	0.008
hexyl acetate	-2.841	-0.077
1,1,2,2-tetrachloroethane	-2.835	-0.012
furfural	-2.822	2.289
E-2-heptenal	-2.802	2.055
1,2-dichlorobenzene	-2.780	0.208
cyclohexanone	-2.730	0.335
3-octanone	-2.697	0.129
octanal	-2.692	1.963
2-ethoxyethyl acetate	-2.637	0.339
tridecane	-2.625	-1.028
indene	-2.584	0.448
3-ethylpyridine	-2.566	0.508
1,3-diisopropylbenzene	-2.557	0.012
2-methylcyclohexanone	-2.543	0.478
propanoic acid	-2.540	2.032
1-hexanol	-2.486	-0.134
benzaldehyde	-2.485	0.734
2-propoxyethanol	-2.472	0.037
1,2,4,5-tetramethylbenzene	-2.459	0.273

Table 12.4 *(Continued)*

VOC	log (1/NPT)	log (1/ODT)
tridecene	-2.413	-0.739
1,4-diisopropylbenzene	-2.410	0.110
nonanal	-2.265	2.274
isobutyric acid	-2.237	2.423
cyclohexanol	-2.218	0.336
2-ethylhexyl acetate	-2.207	0.350
tetradecane	-2.169	-0.930
γ-butyrolactone	-2.168	1.171
1,2,4-trichlorobenzene	-2.114	0.687
decanal	-2.088	2.382
2-methylbenzaldehyde	-2.044	1.081
methyl benzoate	-2.034	0.967
2-butoxyethanol	-2.025	0.493
1-octen-3-ol	-1.947	0.412
acetophenone	-1.925	1.237
2,2-dimethylpropanoic acid	-1.901	2.815
2-ethylhexyl acrylate	-1.827	0.599
2-ethyl-1-hexanol	-1.794	0.464
naphthalene	-1.788	1.236
longifolene	-1.784	-0.103
pentadecane	-1.714	-0.896
pentanoic acid	-1.671	3.087
α-cedrene	-1.664	-0.058
1-octanol	-1.575	0.655
linalyl acetate	-1.538	0.877
1,2-ethanediol	-1.498	0.848
caryophyllene	-1.489	-0.064
hexadecane	-1.259	-0.926
hexanoic acid	-1.258	3.512
1,2-propanediol	-1.244	1.151
octylbenzene	-1.186	0.487

Table 12.4 *(Continued)*

VOC	log (1/NPT)	log (1/ODT)
α-terpineol	-1.033	1.424
benzyl alcohol	-1.019	1.687
phenol	-0.977	1.506
o-cresol	-0.888	1.640
n-methyl-2-pyrrolidone	-0.822	2.599
heptadecane	-0.804	-1.019
heptanoic acid	-0.803	3.929
dimethyl adipate	-0.473	2.436
acenaphthene	-0.367	2.089
2,5-dimethylphenol	-0.361	1.998
2,4-dimethylphenol	-0.359	2.021
octadecane	-0.350	-1.177
octanoic acid	-0.327	4.314
decylbenzene	-0.306	0.468
2-phenoxyethanol	-0.117	2.539
3,5-dimethylphenol	-0.113	2.208
3,4-dimethylphenol	0.032	2.342
1,4-butanediol	0.176	2.514
hexachlorobenzene	0.299	1.708
ε-caprolactam	0.413	3.479
dimethyl phthalate	0.527	3.299
eicosane	0.560	-1.686
1-naphthol	1.503	3.508
2-naphthol	1.683	3.653
hexadecanoic acid	3.209	4.861
max. value =	3.209	4.861
min. value =	-7.000	-5.311

The relative significance of the descriptors in Eqs. (12.2) and (12.5) can be calculated from the product of the coefficient and the mean value of the corresponding descriptor. The percentage weights are then as shown in Table 12.5. As can be observed in this table, the $l \cdot \mathbf{L}$ term is by far the most important. The latter is related to the size of the

solute, and very roughly to the molecular weight. Therefore, the lower molecular weight VOCs are predicted to have small values of log (1/NPT) and log (1/ODT), and hence large NPT and ODT values. The large and very lipophilic VOCs such as octadecane, decyl benzene and ε-caprolactam are predicted to have larger values of log (1/NPT) and log (1/ODT), and hence moderate NPT and ODT values.

Table 12.5
Relative significance of the descriptors, in percentage

	E	**S**	**A**	**B**	**L**
Eq. (12.2)	0	18	11	9	62
Eq. (12.5)	6	9	2	7	75

Our analysis shows that the greater the value of **L** for a given functionally substituted compound, the larger will be log (1/NPT) and log (1/ODT), the smaller will be the NPT and ODT value and the more potent will be the VOC. This is shown in Table 12.6, where the VOCs are listed in order of increasing potency. There is a very good correspondence with **L**, illustrating the importance of this descriptor in Eqs. (12.2) and (12.5).

Table 12.6
The effect of size on log(1/NPT) and log(1/ODT) values for some alcohols

Compound name	L	MW	log (1/NPT)	log (1/ODT)
Ethanol	1.485	46.1	-4.402	-2.629
Propan-2-ol	1.764	60.1	-4.298	-2.368
Propan-1-ol	2.031	60.1	-3.910	-1.879
2-Methylpropan-1-ol	2.413	74.1	-3.616	-1.469
Butan-1-ol	2.601	74.1	-3.396	-1.177
3-Methylbutan-1-ol	3.011	88.2	-3.077	-0.796
2-Methylbutan-1-ol	3.011	88.2	-3.077	-0.789
Pentan-1-ol	3.106	88.2	-2.940	-0.622
Hexan-1-ol	3.610	102.2	-2.486	-0.134
Oct-1-en-3-ol	4.168	128.2	-1.947	0.412
Octan-1-ol	4.619	130.2	-1.575	0.655

By plotting again the histograms of the values of the descriptors in Table 12.4 in the desired range, showing the frequency of the descriptors (Figure 12.7), these can be compared with those descriptors in Tables 12.1 and 12.3 (Figure 12.2 and 12.3) used to obtain the equations for nasal pungency and odour detection threshold. Note that it is only within or slightly this space that Eqs. (12.2) and (12.5) can be considered to be valid and can be used to predict log (1/NPT) and log (1/ODT) values.

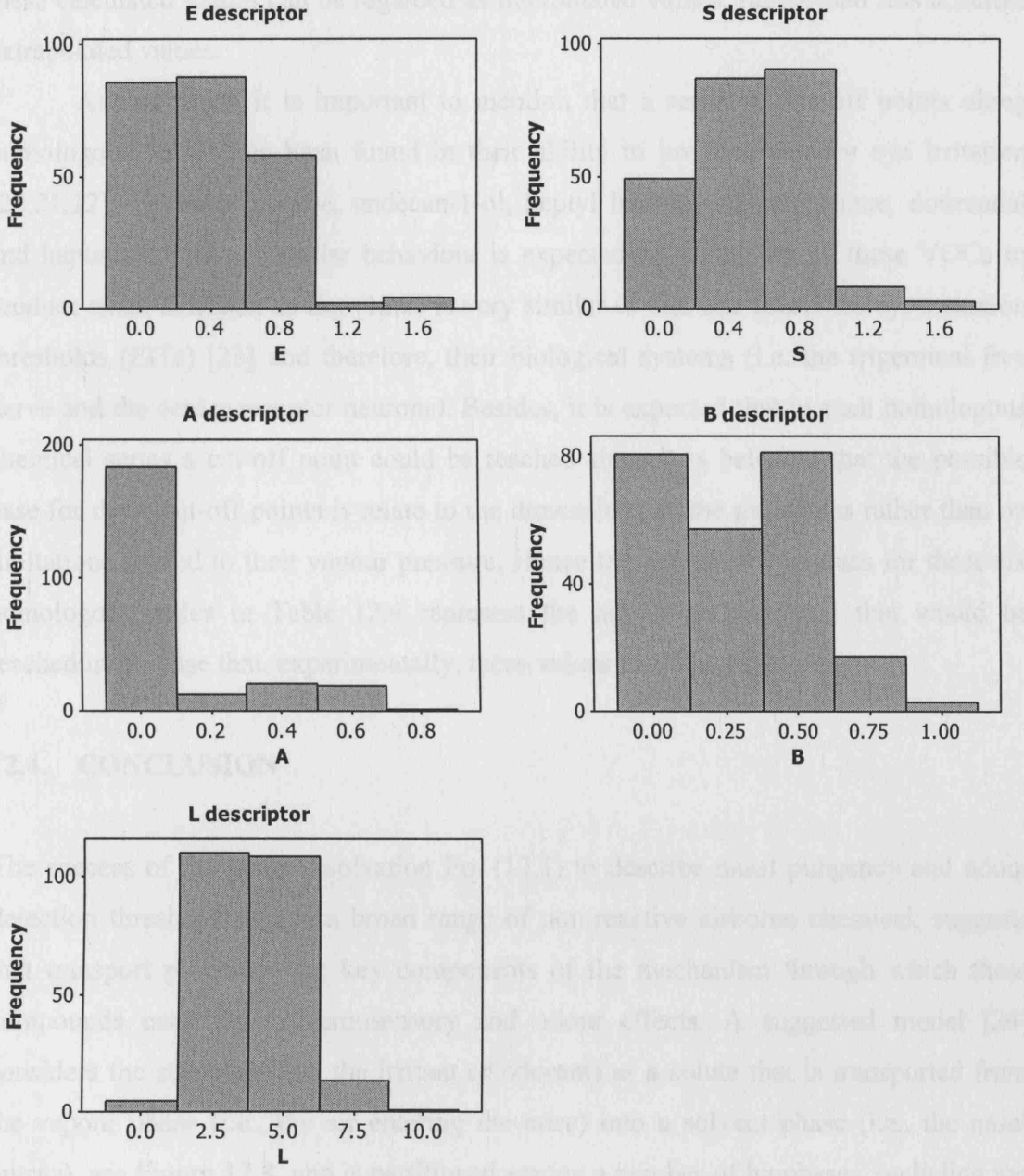


Figure 12.7. Histograms of the descriptors in Table 12.4.

As observed in Figures 12.2 and 12.3, the ranges shown in the histograms look rather small, and they are indeed. However, they cover practically the total range of the

descriptors of those VOCs found in indoor air (Table 12.4) and for which Eqs. (12.2) and (12.5) have been used to predict their nasal pungency and odour detection threshold value. Only a few compounds in Table 12.4 have some of their descriptors above the maximum ranges in Figures 12.2 and 12.3, i.e. for the E, S, A and B descriptors, 0.8, and for L, 7.0. This means that almost all of the predicted log (1/NPT) and log (1/ODT) values in Table 12.4 are calculated from descriptors within the desired range, and thus these calculated values can be regarded as interpolated values, rather than less accurate extrapolated values.

At this point, it is important to mention that a series of cut-off points along homologous VOCs has been found in their ability to produce sensory eye irritation [20,21,22], i.e. decyl acetate, undecan-1-ol, heptyl benzene, 2-tridecanone, dodecanal and heptanoic acid. A similar behaviour is expected in the ability of these VOCs to produce nasal irritation as Eq. (12.2) is very similar to that one found for eye irritation thresholds (EITs) [23] and therefore, their biological systems (i.e. the trigeminal free nerve and the ocular receptor neurons). Besides, it is expected that in each homologous chemical series a cut-off point could be reached since it is believed that the possible base for these cut-off points is relate to the dimensions of the molecules rather than on limitations related to their vapour pressure. Hence the log (1/NPT) values for these six homologous series in Table 12.4 represent the maximum threshold that would be reached in the case that, experimentally, these values could be measured.

12.4. CONCLUSION

The success of the general solvation Eq. (12.1) to describe nasal pungency and odour detection threshold toward a broad range of non-reactive airborne chemical, suggests that transport processes are key components of the mechanism through which these compounds exert their chemosensory and odour effects. A suggested model [24] considers the stimulus (i.e., the irritant or odorant) as a solute that is transported from the vapour phase (i.e., the air entering the nose) into a solvent phase (i.e., the nasal mucus), see Figure 12.8, and is partitioned among a number of biophases, including the one responsible for the biological response (i.e., the free nerve endings of the trigeminal nerve or the olfactory receptor neurons). In other words, the model is said to account for the pungency and odour potency of compounds acting mainly via "transport" (or "transfer") processes and not those acting mainly via very specific "lock-and-key"

stimulus-receptor interactions where well-defined restrictions in size, shape, and spatial configuration play a crucial role.

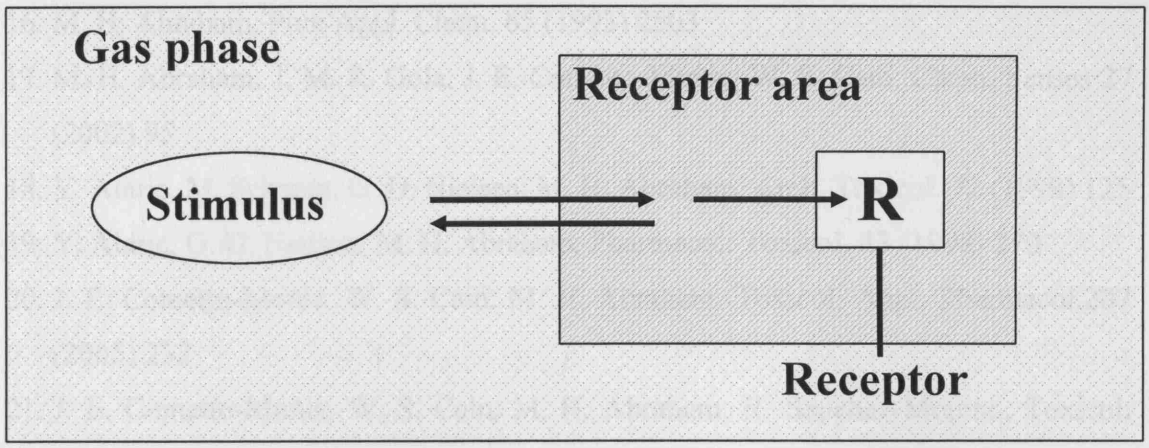


Figure 12.8. The two-stage mechanism of biological activity of VOCs.

12.5. REFERENCES

1. S.K. Brown, M.R. Sim, M.J. Abramson, C.N. Gray, Indoor Air 4 (1994) 123
2. Y. Alarie, Environ. Health Perspect. 42 (1981) 9
3. J. Muller, G. Greff, Food Chem. Toxicol. 22 (1984) 661
4. D. W. Roberts, Chem. Biol. Interact. 57 (1986) 325
5. R. L. Doty, W. E. Brugger, P. C. Jurs, M. A. Orndorff, P. F. Snyder, L. D. Lowry, Physiol. Behav. 20 (1978) 175
6. G. D. Nielsen, E. S. Thomsen, Y. Alarie, Acta Pharmacol. Nord. 1 (1990) 31
7. G. D. Nielsen, L. F. Hansen, Y. Alarie (1992). Irritation of the upper airways. Mechanisms and structure–activity relationships. In Chemical, microbiological, health and comfort aspects of indoor air quality—State of the art in SBS, eds. H.Knöppel and P. Wolkoff, Dordrecht: Kluwer, p. 99
8. K. M. Hau, D. W. Connell, B. J. Richardson, Toxicol. Sci. 47 (1999) 93
9. M. H. Abraham, J. Andonian-Haftvan, J. E. Cometto-Muñiz, W. S. Cain, Fundam. Appl. Toxicol. 31 (1996) 71
10. A. Dravnieks, Ann. NY Acad. Sci. 237 (1974) 144
11. P. Laffort, F. Patte, J. Chromatogr. 406 (1987) 51
12. M. Chastrette, Environ. Res. 6 (1997) 215
13. M. Chastrette, Chem. Senses 23 (1998) 181

14. M. Devos, F. Patte, J. Roualt, P. Laffort, L. J. Van Gemert (1990). Standardized Human Olfactory Thresholds. Oxford, IRL Press
15. T. Yamanaka, Chem. Senses 29 (1995) 471
16. M. H. Abraham, Pure Appl. Chem. 65 (1993) 2503
17. M. H. Abraham, J. M. R. Gola, J. E. Cometto-Muñiz, W. S. Cain, Chem. Senses 27 (2002) 95
18. Y. Alarie, M. Schaper, G. D. Nielsen, M. H. Abraham, Arch. Toxicol. 72 (1998) 125
19. Y. Alarie, G. D. Nielsen, M. H. Abraham, Pharmacol. Toxicol. 83 (1998) 270
20. J. E. Cometto-Muñiz, W. S. Cain, M. H. Abraham, Toxicol. Appl. Pharmacol.207 (2005) 232
21. J. E. Cometto-Muñiz, W. S. Cain, M. H. Abraham, R. Sanchez-Moreno, Toxicol. Sci. 91 (2006) 600
22. J. E. Cometto-Muñiz, W. S. Cain, M. H. Abraham, R. Sanchez-Moreno, Neuroscience 145 (2007) 1130–1137
23. M. H. Abraham, J. M. R. Gola, J. E. Cometto-Muñiz, W. S. Cain, Indoor Built Environ. 10 (2001) 252
24. M. H. Abraham, G. D. Nielsen, Y. Alarie, J. Pharm. Sci. 83 (1994) 680

Chapter 13 A Quantitative Structure-Activity Analysis on the Relative Sensitivity of the Olfactory and the Nasal Trigeminal Chemosensory Systems

13.0. INTRODUCTION

Humans rely principally on two chemosensory systems to detect airborne chemicals: olfaction and chemesthesis. The sense of smell is restricted to the nasal cavity and mediated by the olfactory nerve. In contrast, chemesthesis [1], or chemical feel, is present in all mucosae, also in the skin under the epidermis [2], and is mediated by a variety of nerves, depending on the location of stimulation. Due to their direct exposure to the air we breathe and that surrounds us, the nasal and the ocular mucosa are common sites of chemesthetic stimulation [3]. Nasal chemesthesis includes sensations such as stinging, freshness, prickling, piquancy, tingling, irritation, burning and the like, which, due to their sharp nature, we group together under the term nasal pungency. Chemesthesis in both sites is mediated by the trigeminal nerve (see review in Doty and Cometto-Muñiz [4]). In the present work we will focus on the comparative sensitivity of olfaction and nasal chemesthesis towards volatile organic compounds (VOCs), using a quantitative structure-activity relationship (QSAR) approach.

Odorants reaching the lumen above the olfactory epithelium transfer from the gas phase into the mucus phase and continue to be distributed among the various biophases until they reach the olfactory receptors (ORs) [5] in the membrane of the olfactory sensory neurons (OSNs) cilia [6,7]. ORs belong to the large family of G-protein coupled receptors (GPCRs) [8,9]. In humans there are about 388 genes coding for functional ORs and about 414 pseudogenes that do not code for functional ORs [10]. Each odorant is believed to activate a specific pattern of ORs [11].

Irritants entering the nasal cavity also transfer from the gas phase into the mucus and other biophases until they reach chemesthetic receptors within free nerve endings of the trigeminal nerve [12], particularly from C and A_{delta} fibers. Trigeminal chemoreceptors include vanilloid [13,14], nicotinic acetylcholine [15,16,17], and menthol [18,19] receptors. Capsaicin, menthol, and a variety of pungent compounds stimulate sensory nerve fibers via activation of members of a family of transient receptor potential (TRP) channels [20,21,22,23] that includes about 30 members

[24,25]. These and other receptors and mechanisms [26], including cell damage by reactive VOCs and consequent release of nociception mediators [27], could play a role in nasal chemesthesis as evoked by common VOCs, including alcohols, esters, ketones, alkylbenzenes, aldehydes, etc. [28].

A previous separate QSAR analysis on nasal pungency thresholds (NPTs) [29] and odor detection thresholds (ODTs) [30] revealed that "selective" processes (i.e., chemically-broad, transfer driven effects) overwhelmingly dominate chemesthetic detection, whereas both selective and "specific" (i.e., chemically-narrow, ligand-receptor interactions) processes control olfactory potency. To understand further the nature of the chemical factors that influence ODT values, we have explored here the possibility that NPT values could be used to estimate *selective* effects in ODTs, thus producing a tool to assess the weight of the remaining *specific* (VOC-receptor) effects. The topic opens the window to ponder why certain chemical families or particular compounds (and which ones) could have driven the need for a more specialized and sensitive chemoreception in humans. The present study involves data on 64 VOCs from various chemical series. The compounds are listed in our previous separate QSAR analysis of odor [30] and nasal pungency [29] thresholds. However, an updated list is given in the appendix, Section 13.5.

13.1. MATERIALS AND METHODS

Both odor and nasal pungency involve the transfer of a compound, for example a VOC, from an air stream through a mucus layer into a receptor or receptor area. This environment is likely to be inhomogeneous, being partly a hydrophobic lipid-like area and partly a hydrophilic aqueous-like area. An equation, Eq. (13.1), that seems to be very satisfactory for the correlation and explanation of the transfer of VOCs from the gaseous phase to a large number of solvents or other condensed phases, including biophases, has already been developed [31,32,33,34].

$$SP = c + e \cdot \mathbf{E} + s \cdot \mathbf{S} + a \cdot \mathbf{A} + b \cdot \mathbf{B} + l \cdot \mathbf{L} \qquad (13.1)$$

In Eq. (13.1), **E**, **S**, **A**, **B**, and **L** are properties, or descriptors, of the VOC, and c, e, s, a, b, and l are regression coefficients, as described in detail previously [35]. Briefly, **E** is the excess molar refraction, **S** is the dipolarity/polarizability, **A** and **B** are the overall or

effective hydrogen bond acidity and basicity, respectively, of the VOC, and **L** (log L^{16}) is defined through L^{16}, the VOC gas-hexadecane partition coefficient at 298 K, which is a measure of the lipophilicity of the VOC. In turn, the regression coefficients are not merely fitted coefficients since they define the complementary physicochemical properties that characterize the receptor environment or biophase most receptive to the VOC [36]. SP is either a physicochemical property of a VOC, such as log K where K is the gas to solvent partition coefficient for a series of VOCs into a given solvent or condensed phase; or a biological property of a VOC, such as an odor or nasal pungency threshold for a series of VOCs [37].

When Eq. (13.1) was applied to NPTs, as log(1/NPT), a very good correlation that accounted for more than 95 % of the total effect was obtained [29]. This strongly suggests that the factors that influence NPTs are those that influence the transfer of VOCs from the gas phase to condensed phases, that is from the gas phase to the receptor phase, and that VOC-receptor interactions are of secondary importance. However, when Eq. (13.1) was applied to ODTs, as log(1/ODT), a much poorer correlation was found [30]. Only by excluding families of compound such as the aldehydes and carboxylic acids or by assigning a special descriptor to these families could a satisfactory correlation be obtained. This suggests that ODT values are partly influenced by transfer from the gas phase to the receptor phase, and must partly be influenced by VOC-receptor interactions. We can further suggest that the transfer-type influences are selective, in that different VOCs are transported from the gas phase to condensed phases with different equilibrium constants, but that VOC-receptor interactions are more specific, in that VOCs with similar structural features may have considerably different affinities for the receptor or effectiveness to stimulate it.

In order to obtain more information on the factors that influence ODT values, we now explore the possibility that NPT values could be used to estimate the selective factors, that is to separate out the selective transport related effects and to leave only the specific (VOC-receptor) effects. The present study involves data on 64 VOCs from various chemical series. The compounds are listed in a previous separate QSAR analysis of odor [30] and nasal pungency [29] thresholds, carried out by Abraham and co-workers.

13.2. RESULTS AND DISCUSSION

ODTs and NPTs were correlated against Abraham's descriptors, using Eq. (13.1). The aim was to obtain a similar equation for both sets of data to make possible a comparison between them. To do so, compounds that were outliers in the equation for ODT, that is aldehydes and carboxylic acids, were left out in both cases. In addition four compounds that were outliers in the ODT equation and which might act through specific effects were omitted, viz: propanone, methyl acetate, t-butyl acetate and 1-octanol. Only the Abraham descriptors from Eq. (13.1) were used, without including any extra descriptor, or indicator variable. The resulting equations are:

$$\log(1/ODT) = -5.27 + 0.51\,\mathbf{E} + 1.96\,\mathbf{S} + 1.48\,\mathbf{A} + 1.53\,\mathbf{B} + 0.723\,\mathbf{L} \qquad (13.2)$$

$N = 50,\ R^2 = 0.780,\ SD = 0.57,\ F = 31.2$

$$\log(1/NPT) = -7.89 + 0.20\,\mathbf{E} + 1.32\,\mathbf{S} + 2.71\,\mathbf{A} + 1.52\,\mathbf{B} + 0.823\,\mathbf{L} \qquad (13.3)$$

$N = 38,\ R^2 = 0.916,\ SD = 0.28,\ F = 70.1$

As can be observed, all the coefficients, except the **A**-coefficient, are quite similar in Eq. (13.2) and Eq. (13.3). Hence NPT values are more affected by VOC hydrogen bond acidity than are the ODT values. The number of NPT values (N=38) is considerably less than the number of ODT values (N=50). In order to obtain a compound-by-compound comparison for all the ODT values, it was therefore decided to use Eq. (13.3) to calculate log(1/NPT) values for all the VOCs for which we had ODT values. These will be referred to as Clog(1/NPT) values (where "C" stands for "calculated").

The observed values of log(1/ODT) were then regressed against Clog(1/NPT) for the VOCs that are likely to act by selective effects only, and Eq. (13.4) was obtained.

$$\text{Log}\,(1/ODT) = 2.321 + 0.939\,\text{Clog}(1/NPT) \qquad (13.4)$$

$N = 50,\ R^2 = 0.747,\ SD = 0.58,\ F = 141.7$

If the **A**-descriptor is used as another independent variable, the regression improves slightly, as shown in Eq. (13.5).

Log (1/ODT) = 2.430 + 0.943 Clog(1/NPT) - 0.955 **A** (13.5)

N = 50, R^2 = 0.764, SD = 0.57, F = 76.2

Both Eq. (13.4) and Eq (13.5) reproduce the values of log (1/ODT) as well as does the full Eq. (13.2), for selectively acting VOCs. Either equation can then be used as a 'base line' for selective effects, and can be used to identify compounds that yield odor detection thresholds through a combination of selective and specific effects. This is illustrated in Figure 13.1, where is plotted log (1/ODT) values for the 50 VOCs used in Eqs. (13.4) and (13.5), plus aldehydes and acids, against calculated values from Eq. (13.5).

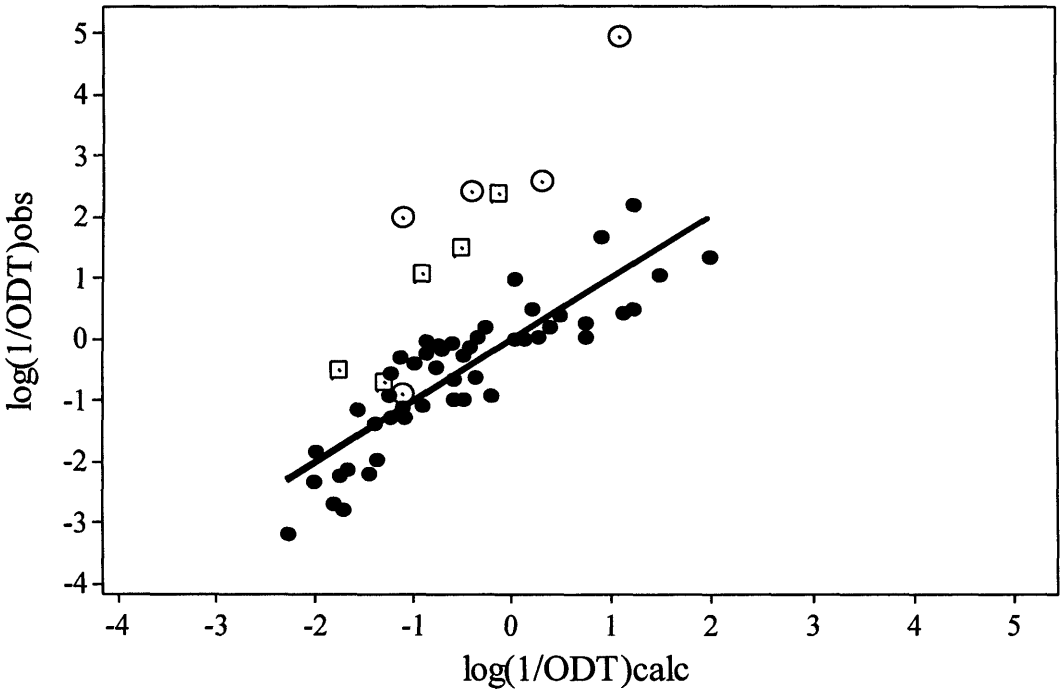


Figure 13.1. A plot of log(1/ODT)obs against log(1/ODT) calculated on Eq. (13.5), showing the specific effects of aldehydes □ and acids ○. The regression line is for the selective compounds.

The five aldehydes are more potent than calculated by an average of 1.7 log units, and the acids (excluding formic acid) by an average of 3.0 log units. These are present estimates of the specific effect of the two series of VOCs. In a previous analysis of odor detection thresholds [30] aldehydes and acids were accommodated into general equations by adding an indicator variable that increased the potency of these compounds by 1.6 or 2.0 log units, depending on the exact form of the equation; the increased potency was for an average for the aldehydes and acids taken together. The present

results for aldehydes and acids taken separately is in line with the previous analysis, but it is suggested that the present method affords a much better estimate of the 'specific' effect on odor detection thresholds. It can be seen that the aldehydes and acids provoke ODT through extra specific as well as selective effects because data are available for several other series for which deviations from Eq. (13.5) can be calculated.

In Table 13.1 are given the values of the average error (AE), the absolute average error (AAE) and the standard deviation (SD) of the observed and calculated values for the various series. The key statistics are AE and AAE. The average error denotes deviation from Eq. (13.5), either in a positive or negative sense {AE = (calculated –observed values)/N where N is the number of data points)}. If AE and AAE are compared, it is then possible to deduce whether a given value of AE is due to random deviations or systematic deviations from Eq. (13.5).

Table 13.1
Deviations from Eq. (13.5) for various series

Series	N	AE	AAE	SD
Alcohols	11	-0.06	0.69	0.77
Acetates	10	0.04	0.43	0.53
Ketones	3	0.11	0.26	0.33
Alkyl benzenes	8	-0.27	0.35	0.45
Aldehydes	5	1.70	1.70	2.05
Acids	4	3.04	3.04	3.57

For the first four series, the AE values are very small, so that there are no systematic deviations. The numerically larger AAE values then represent random deviations, as do the corresponding SD values. These range from 0.33 to 0.77 log units in accord with the SD value of 0.57 log units in Eq. (13.5). However, for the aldehydes and acids, AE is identical to AAE - all the deviations are of the same sign and are then systematic and not random. The SD value for the aldehydes is 3.6 times the SD in Eq. (13.5) and for the acids is 6.3 times the SD, so that the deviations from Eq. (13.5) are very large indeed. It is these systematic, not random, deviations from Eq. (13.5) that lead to the conclusion that aldehydes and acids exert effects on ODTs through a combination of specific and selective effects. For individual VOCs the position is not that straightforward. Of the four outliers that have been identified for ODTs, 1-octanol (2.15 obs, 0.41 calc) and

tert-butyl acetate (-0.11 obs, -1.49 calc) are more potent than calculated through Eq. (13.5), but whether these effects are really due to specific VOC-receptor interactions, rather than error in either the ODT determinations or the descriptors is difficult to assess. Methyl acetate (-3.46 obs, -2.06 calc) and propanone (-4.07 obs, -2.02 calc) are both much less potent than calculated through Eq. (13.5); this cannot be due to any (extra) specific effect at all, and suggests that there are still some factors that still need to be accounted for.

Now that Clog(1/NPT) values have been used to determined the specific effect of aldehydes and acids on ODT values, the log(1/NPT) values themselves can now be used in order to obtain a correlation between observed ODT values and observed NPT values for compounds that exert their influence through selective effects.

$$\text{Log}(1/\text{ODT}) = 3.562 + 1.282\ \log(1/\text{NPT}) \qquad (13.6)$$

$N = 34,\ R^2 = 0.819,\ SD = 0.49,\ F = 144.5$

$$\text{Log}\ (1/\text{ODT}) = 3.697 + 1.267\ \log(1/\text{NPT}) - 1.457\ \mathbf{A} \qquad (13.7)$$

$N = 34,\ R^2 = 0.867,\ SD = 0.42,\ F = 101.1$

Eq. (13.6) can be used to predict further values of log (1/ODT) for VOCs that are thought to act through selective effects only; the SD value of only 0.49 log units is less than that in the full Eq. (13.2), although the latter is for 50 compounds. Unlike Eq. (13.4) and Eq. (13.5), there is now a substantial gain in goodness of fit if the **A**-descriptor is used as an independent variable, with the SD reduced to 0.42 log units. Hence Eq. (13.7) represents an even better predictive method (see Figure13.2). Descriptor values are known for several thousand compounds [38] and **A**-values are available for the prediction of log (1/ODT) values through Eq. (13.7) for numerous other VOCs. If not, it is possible to calculate **A**-values just from the structure of VOCs [38], so that in most cases Eq. (13.7) can be used for predictions. If an **A**-value is neither known nor is available, then Eq. (13.6) still represents a very good method for the prediction of further values. The quantitative relationship that has been established between ODTs and NPTs for VOCs that act mainly via selective effects can facilitate the identification of outlying odorants for whom additional specific effects play a substantial role in their olfactory detection. These odorants could become prime candidates in the search of the best ligands for orphan ORs [39]. In addition, knowing

the identity of these particularly powerful odorants can provide clues on the evolutionary factors that have driven the sense of smell to carve an enhanced sensitivity towards them [10].

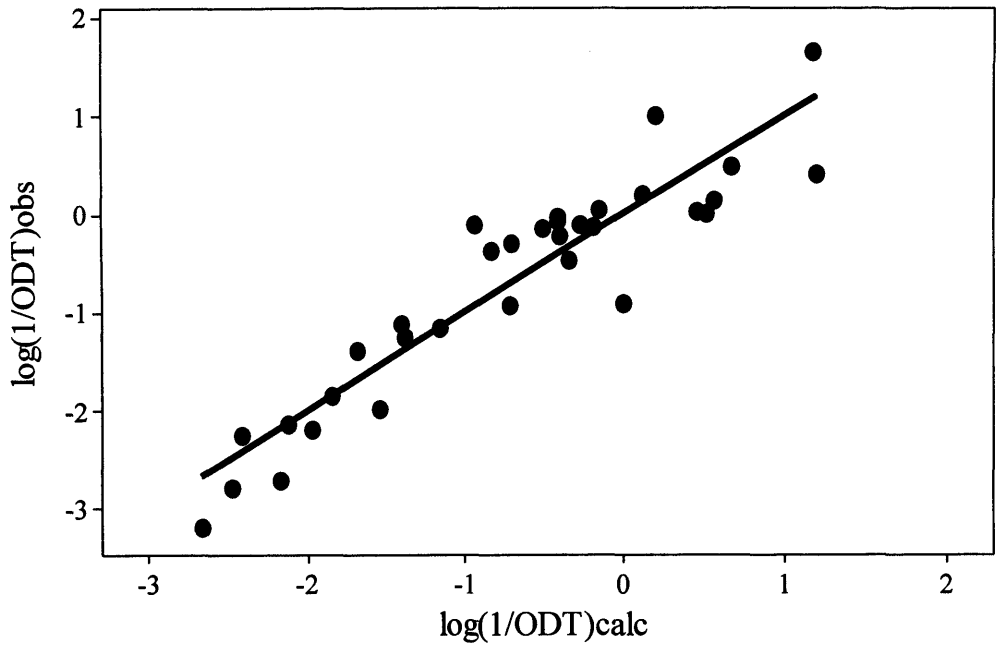


Figure 13.2. A plot of log(1/ODT)obs against log(1ODT)calc on Eq (13.7)

Finally, the reasons why nasal pungency thresholds can be used as a base line for selective effects in another biological end point altogether, can be explored. Possible biological mechanisms of action of VOCs have been set out [40] and Figure 13.3 is shown the 'two-stage' mechanism that was suggested [40]. In the first stage the VOC is transferred from the vapor phase to a receptor phase or receptor area, and in the second stage the VOC activates a receptor. Now transfer from the vapor phase to typical organic phases involves selective effects, not specific effects. Thus if the main step in the mechanism of nasal pungency is the first stage, this would account for structural effects in the VOCs being selective only. Some information on this can be obtained by comparing the coefficients in the selective NPT and ODT equations, Eq. (13.2) and Eq. (13.3), with those for transfer from the gas phase to various solvents [41,42] and biological phases [33,43,44,32,33,34]. Details are in Table 13.2. The coefficients in these equations reflect the chemical properties of the phases, so that the nearer is one set of coefficients to another set, the closer are the chemical properties of the phases. It is rather difficult to assess the closeness of sets of coefficients just by eye, and it is convenient to use principle components analysis (PCA). In this method, the five

columns of coefficients, *e* to *l*, are transformed into five orthogonal columns of data (five PCs) that contain the same information as the original columns. The first two PCs contain 80% of the total information, in the present case, and so a plot of the scores of PC2 against PC1 will provide a visual estimate of how close are the sets of coefficients. The coefficients for the phases investigated are in Table 13.2, and the PC plot is shown as Figure 13.4.

Table 13.2
Coefficients in equations for gas to solvent or phase transfer

Solvent or phase	No	c	e	s	a	b	l
log(1/NPT)	1	-7.89	0.20	1.32	2.71	1.52	0.823
log(1/ODT)	2	-5.27	0.51	1.96	1.48	1.53	0.723
Blood, 37°C	3	-1.07	0.46	1.08	3.74	2.58	0.376
Brain, 37 °C	4	-0.99	0.26	0.41	3.36	2.03	0.591
Muscle, 37 °C	5	-1.04	0.21	0.72	3.24	2.47	0.463
Liver, 37 °C	6	-0.92	0.08	0.77	2.79	2.09	0.560
Fat, 37 °C	7	-0.05	0.05	0.73	1.78	0.33	0.743
Olive oil 37 °C	8	-0.16	-0.25	0.86	1.66	0.00	0.873
1-Octanol	9	-0.20	0.00	0.71	3.52	1.43	0.858
Methanol (dry)	10	0.00	-0.22	1.17	3.70	1.43	0.769
Ethanol (dry)	11	0.01	-0.21	0.79	3.64	1.31	0.853
1-Butanol (dry)	12	-0.04	-0.28	0.54	3.78	1.00	0.934
1-Octanol (dry)	13	-0.12	-0.20	0.56	2.56	0.70	0.939
N-Methylformamide(dry)	14	-0.60	-0.26	2.00	4.56	0.43	0.706
Ethyl acetate (dry)	15	0.20	-0.34	1.25	2.95	0.00	0.917
Acetone (dry)	16	0.15	-0.28	1.52	3.26	0.00	0.863
Ether (dry)	17	0.21	-0.17	0.87	3.40	0.00	0.882
Acetonitrile (dry)	18	-0.01	-0.60	2.46	2.09	0.42	0.738
Chloroform	19	0.12	-0.47	1.20	0.14	1.43	0.994
Ethylene glycol (dry)	20	-0.90	0.22	1.43	4.47	2.69	0.568
Hexadecane	21	0.00	0.00	0.00	0.00	0.00	1.000
Cyclohexane	22	0.16	-0.11	0.00	0.00	0.00	1.013
Toluene	23	0.12	-0.22	0.94	0.47	0.10	1.012

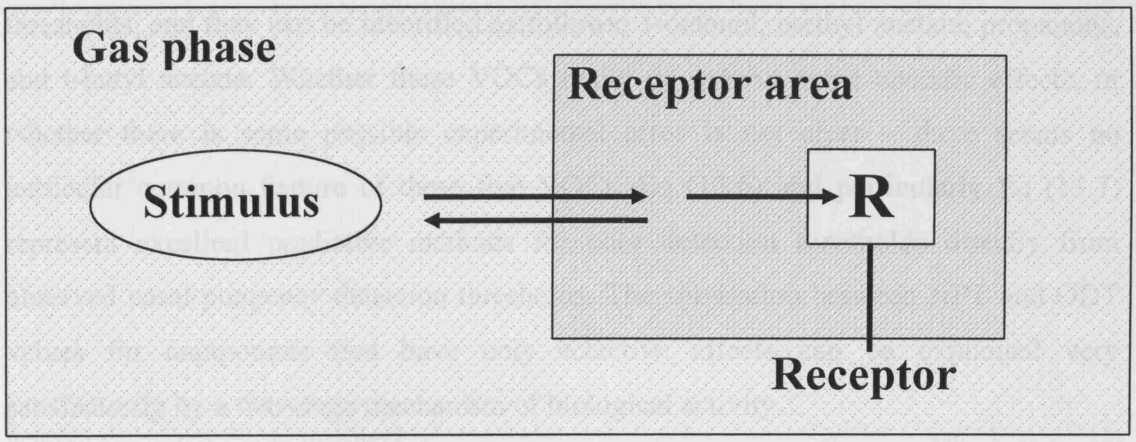


Figure 13.3. The two-stage mechanism of biological activity of VOCs.

The coefficients for the NPT and ODT equations are quite near to each other and to coefficients for transfer from the gas phase to biological phases (brain, muscle, liver) and organic solvents (wet 1-octanol, methanol, ethanol). All these solvents or phases have substantial values of the a- and b-coefficients. Transfer to all the solvents and biological phases shown in Table 13.3 involves selective, not specific, structural effects of the VOCs. Hence if the main step in a mechanism involves stage 1, or if only VOCs that act by selective effects are included, it is to be expected that the coefficients in the corresponding equations will be close to some particular solvent or biological phase. This is exactly what is found for the NPT or ODT equations. Of course the mechanism of nasal pungency, or odor, detection thresholds will involve VOCs passing from the gas phase through various layers of materials to the receptor phase [5]. But in an equilibrium situation, the overall equilibrium constant depends only on the concentrations in the initial phase (the gas phase) and the final phase (the receptor phase) – the intermediate phases in this context are irrelevant.

13.3. CONCLUSION

It is possible to separate out specific effects from selective effects in odor detection thresholds by the use of nasal pungency thresholds used as an indication of selective chemosensory effects. The main VOCs that show specific effects are the aldehydes and carboxylic acids, except for formic acid. Although this VOC appears 'normal' it is possible that this is a fortuitous combination of more than one specific effect. There are other VOCs that appear also to exhibit some specific effects as regards odour

thresholds, and they can be identified as follows: 1-octanol, methyl acetate, propanone, and t-butyl acetate. Whether these VOCs really do exhibit some specific effects, or whether there is some possible experimental error is not clear – there seems no particular common feature of these five VOCs. Eq (13.6) and particularly Eq (13.7) represent excellent predictive methods for odor detection thresholds directly from observed nasal pungency detection thresholds. The correlation between NPT and ODT values for compounds that have only selective effects can be explained very satisfactorily by a two-stage mechanism of biological activity.

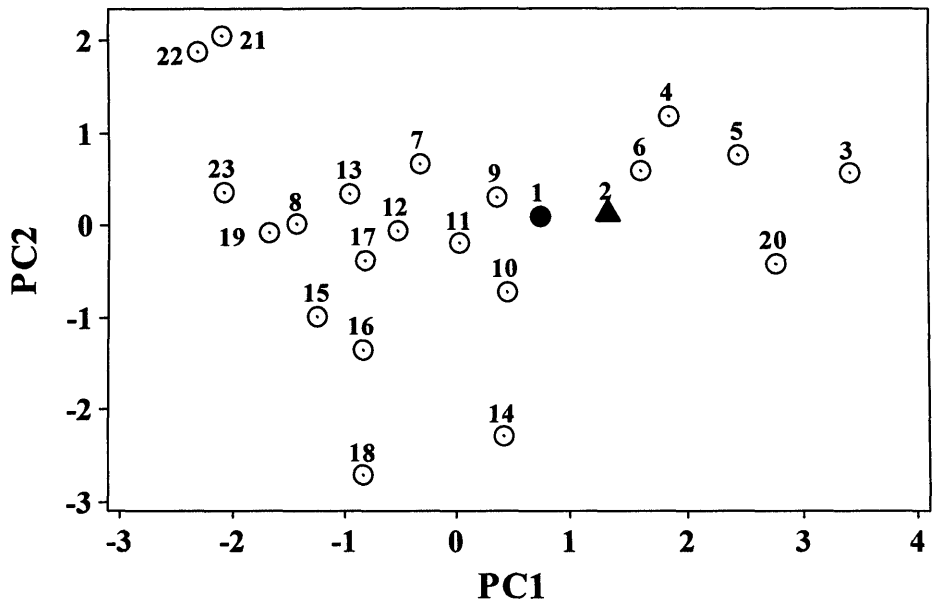


Figure 13.4. A plot of the scores of PC2 against the scores of PC1. Points numbered as in Table 13.2: • NPT, ▲ODT.

13.4. REFERENCES

1. B. Bryant, W. L. Silver, (2000) Chemesthesis: The Common Chemical Sense. In T. E. Finger, W. L. Silver, and D. Restrepo, (ed.) *The Neurobiology of Taste and Smell.* 2nd Edition. Wiley-Liss, New York, pp. 73-100
2. C. A. Keele, Arch. Int. Pharmacodyn. Ther. 139 (1962) 547
3. R. L. Doty, J. E. Cometto-Muñiz, A. A. Jalowayski, P. Dalton, M. Kendal-Reed, M. Hodgson, Crit. Rev. Toxicol. 34 (2004) 85

4. R. L. Doty, J. E. Cometto-Muñiz, (2003) Trigeminal chemosensation. In Doty, R.L. (ed.) *Handbook of Olfaction and Gustation*, 2nd Edition. Marcel Dekker, New York, pp. 981-1000
5. N. E. Rawson, K. K. Yee, Adv. Otorhinolaryngol. 63 (2006) 23
6. R. J. Flannery, D. A. French, S. J. Kleene, Biophys. J. 91 (2006) 179
7. K. Schwarzenbacher, J. Fleischer, H. Breer, Histochem. Cell Biol. 123 (2005) 419
8. H. Breer, Anal. Bioanal. Chem. 377 (2003) 427
9. E. R. Liman, Pflugers Arch. 453 (2006) 125
10. Y. Niimura, M. Nei, J. Hum. Genet. 51 (2006) 505
11. B. Malnic, J. Hirono, T. Sato, L. B. Buck, Cell 96 (1999) 713
12. T. E. Finger, V. L. St Jeor, J. C. Kinnamon, W. L. Silver, J. Comp. Neurol. 294 (1990) 293
13. W. L. Silver, T. R. Clapp, L. M. Stone, S. C. Kinnamon, Chem. Senses 31 (2006) 807
14. M. Tominaga, T. Tominaga, Pflugers Arch. 451 (2005) 143
15. H. Alimohammadi, W. L. Silver, Chem. Senses 25 (2000) 61
16. N. Thuerauf, M. Kaegler, R. Dietz, A. Barocka, G. Kobal, Psychopharmacology (Berl.) 142 (1999) 236
17. N. Thuerauf, K. Markovic, G. Braun, S. Bleich, U. Reulbach, J. Kornhuber, J. Lunkenheimer, Neuropsychopharmacology 31 (2006) 450
18. N. Damann, M. Rothermel, B. G. Klupp, T. C. Mettenleiter, H. Hatt, C. H. Wetzel, BMC Neurosci. 7 (2006) 46
19. K. Kobayashi, T. Fukuoka, K. Obata, H. Yamanaka, Y. Dai, A. Tokunaga, K. Noguchi, J. Comp. Neurol. 493 (2005) 596
20. S. E. Jordt, D. M. Bautista, H. H. Chuang, D. D. McKemy, P. M. Zygmunt, E. D. Hogestatt, I. D. Meng, D. Julius, Nature 427 (2004) 260
21. L. J. Macpherson, B. H. Geierstanger, V. Viswanath, M. Bandell, S. R. Eid, S. Hwang, A. Patapoutian, Curr. Biol. 15 (2005) 929
22. L. J. Macpherson, S. W. Hwang, T. Miyamoto, A. E. Dubin, A. Patapoutian, G. M. Story, Mol. Cell Neurosci. 32 (2006) 335
23. M. Trevisani, D. Smart, M. J. Gunthorpe, M. Tognetto, M. Barbieri, B. Campi, S. Amadesi, J. Gray, J. C. Jerman, S. J. Brough, D. Owen, G. D. Smith, A. D. Randall, S. Harrison, A. Bianchi, J. B. Davis, P. Geppetti, Nat. Neurosci. 5 (2002) 546
24. C. Montell, Sci. STKE 2005 (2005) re3

25. I. S. Ramsey, M. Delling, D. E. Clapham, Annu. Rev. Physiol. 68 (2006) 619
26. T. Inoue B. P. Bryant, Pain 117 (2005) 193
27. S. P. Sutherland, S. P. Cook, E. W. McCleskey, Prog. Brain Res. 129 (2000) 21
28. J. E. Cometto-Muñiz, (2001) Physicochemical basis for odor and irritation potency of VOCs. In Spengler, J.D., Samet, J. and McCarthy, J.F. (ed.) *Indoor Air Quality Handbook.* McGraw-Hill, New York, pp. 20.1—20.21.
29. M. H. Abraham, R. Kumarsingh, J. E. Cometto-Muñiz, W. S. Cain, Arch. Toxicol. 72 (1998) 227
30. M. H. Abraham, J. M. R. Gola, J. E. Cometto-Muñiz, W. S. Cain, Chem. Senses 27 (2002) 95
31. M. H. Abraham, Chem. Soc. Rev. 22 (1993) 73
32. M. H. Abraham, A. Ibrahim, W. E. Acree, Jr., Eur. J. Med. Chem. 41 (2006) 494
33. M. H. Abraham, A. Ibrahim, W. E. Acree, Jr., Chem. Res. Toxicol. 19 (2006) 801
34. M. H. Abraham, A. Ibrahim, W. E. Acree, Jr., Eur. J. Med. Chem. (2007)
35. M. H. Abraham, A. Ibrahim, A. M. Zissimos, J. Chromatogr. A 1037 (2004) 29
36. M. H. Abraham, (1996) The potency of gases and vapors: QSARs — Anesthesia, sensory irritation, and odor. In R. B. Gammage and B. A. Berven (ed.) *Indoor Air and Human Health.* 2nd Edition. CRC Lewis Publishers, Boca Raton, pp. 67-91.
37. M. H. Abraham, J. M. R. Gola, J. E. Cometto-Muñiz, W. S. Cain, Indoor Built Environ. 10 (2001) 252
38. PharmaAlgorithms, Version 3.0. (2006) ADME Boxes, 591, Indian Road, Toronto, ON M6P 2C4, Canada.
39. A. Mizrahi, H. Matsunami, L. C. Katz, Neuropharmacology 47 (2004) 661
40. M. H. Abraham, G. D. Nielsen, Y. Alarie, J. Pharm. Sci. 83 (1994) 680
41. M. H. Abraham, A. Ibrahim, J. Chem. Inf. Model. 46 (2006) 1735
42. K. R. Hoover, R. Coaxum, E. Pustejovsky, D. M. Stovall, W. E. Acree, M. H. Abraham, Phys. Chem. Liq. 42 (2004) 339
43. M. H. Abraham, A. Ibrahim, Eur. J. Med. Chem. 41 (2006) 1430
44. M. H. Abraham, A. Ibrahim, W. E. Acree, Jr., Chem. Res. Toxicol. 18 (2005) 904

13.5. APPENDIX

Table 13.3
Compounds studied, their descriptors and values of log(1/ODT) and log(1/NPT)

Compound Name	E	S	A	B	L	log(1/ODT)	log(1/NPT)
methanol	0.278	0.44	0.43	0.47	0.970	-3.18	-4.54
ethanol	0.246	0.42	0.37	0.48	1.485	-1.85	-3.95
1-propanol	0.236	0.42	0.37	0.48	2.031	-1.15	-3.40
2-propanol	0.212	0.36	0.33	0.56	1.764	-2.70	-4.26
1-butanol	0.224	0.42	0.37	0.48	2.601	-0.30	-3.04
2-butanol	0.217	0.36	0.33	0.56	2.338	-1.98	-3.76
tert-butyl alcohol	0.180	0.30	0.31	0.60	1.963	-2.78	-4.52
1-pentanol	0.219	0.42	0.37	0.48	3.106	-0.11	-3.23
1-hexanol	0.210	0.42	0.37	0.48	3.610	0.05	-2.60
1-heptanol	0.211	0.42	0.37	0.48	4.115	1.00	-2.32
4-heptanol	0.180	0.36	0.33	0.56	3.850	-0.91	-2.53
1-octanol	0.199	0.42	0.37	0.48	4.619	2.15	-1.85
methyl acetate	0.142	0.64	0.00	0.45	1.911	-3.46	-5.05
ethyl acetate	0.106	0.62	0.00	0.45	2.314	-2.24	-4.83
propyl acetate	0.092	0.60	0.00	0.45	2.819	-1.39	-4.24
butyl acetate	0.071	0.60	0.00	0.45	3.353	-0.38	-3.56
sec-butyl acetate	0.044	0.57	0.00	0.47	3.054	-0.57	-3.50
tert-butyl acetate	0.025	0.54	0.00	0.47	2.802	-0.11	-3.98
pentyl acetate	0.067	0.60	0.00	0.45	3.844	-0.07	-3.22
hexyl acetate	0.056	0.60	0.00	0.45	4.290	0.20	-2.80
heptyl acetate	0.050	0.60	0.00	0.45	4.796	0.01	-2.49
octyl acetate	0.046	0.60	0.00	0.45	5.270	0.41	-1.95
decyl acetate	0.041	0.60	0.00	0.45	6.240	0.50	
dodecyl acetate	0.038	0.60	0.00	0.45	7.219	1.36	
2-propanone	0.179	0.70	0.04	0.49	1.696	-4.07	-5.12
2-pentanone	0.143	0.68	0.00	0.51	2.755	-0.93	-3.47
2-heptanone	0.123	0.68	0.00	0.51	3.760	-0.27	-2.91
2-nonanone	0.113	0.68	0.00	0.51	4.735	0.03	-2.53
toluene	0.601	0.52	0.00	0.14	3.325	-2.19	-4.47

Table 13.3 *(continued)*

Compound Name	E	S	A	B	L	log(1/ODT)	log(1/NPT)
ethyl benzene	0.613	0.51	0.00	0.15	3.778	-1.26	-4.00
propyl benzene	0.604	0.50	0.00	0.15	4.230	-0.47	-3.17
butyl benzene	0.600	0.51	0.00	0.15	4.730	-0.63	
pentyl benzene	0.594	0.51	0.00	0.15	5.230	0.00	
hexyl benzene	0.591	0.50	0.00	0.15	5.720	0.19	
heptyl benzene	0.577	0.48	0.00	0.15	6.219	0.25	
octyl benzene	0.579	0.48	0.00	0.15	6.714	0.43	
butanal	0.187	0.65	0.00	0.45	2.270	-0.48	-4.77
pentanal	0.163	0.65	0.00	0.45	2.851	-0.70	-4.57
hexanal	0.146	0.65	0.00	0.45	3.357	1.10	-3.70
heptanal	0.140	0.65	0.00	0.45	3.865	1.52	-3.13
octanal	0.160	0.65	0.00	0.45	4.361	2.40	-3.24
formic acid	0.343	0.75	0.76	0.33	1.545	-0.89	-2.50
acetic acid	0.265	0.64	0.62	0.44	1.816	2.00	-1.62
butanoic acid	0.210	0.64	0.61	0.45	2.750	2.44	-1.79
hexanoic acid	0.174	0.63	0.62	0.44	3.697	2.59	-1.30
octanoic acid	0.150	0.65	0.62	0.45	4.680	4.96	
cumene	0.602	0.49	0.00	0.16	4.084	-0.03	-3.22
p-cymene	0.607	0.49	0.00	0.19	4.590	-0.12	-3.05
Δ-3-carene	0.511	0.22	0.00	0.10	4.649	-0.22	-3.21
linalool	0.398	0.55	0.20	0.67	4.794	0.02	-2.55
1,8-cineole	0.383	0.33	0.00	0.76	4.688	0.49	-2.37
geraniol	0.513	0.54	0.35	0.63	5.510	1.05	
α-terpinene	0.526	0.25	0.00	0.15	4.715	-0.15	-3.30
γ-terpinene	0.497	0.32	0.00	0.20	4.815	-0.99	
α-pinene	0.446	0.14	0.00	0.12	4.308	-1.28	
β-pinene	0.530	0.24	0.00	0.19	4.394	-1.07	
(r)-(+)- limonene	0.488	0.28	0.00	0.21	4.725	-0.99	
(s)- (-)-limonene	0.488	0.28	0.00	0.21	4.725	-0.66	
β-phenyl ethyl alcohol	0.811	0.86	0.31	0.65	4.628	2.19	
pyridine	0.631	0.84	0.00	0.52	3.022	-0.11	-3.11

Table 13.3 *(continued)*

Compound Name	E	S	A	B	L	log(1/ODT)	log(1/NPT)
menthol	0.400	0.50	0.23	0.58	5.177	1.66	-1.71
1-octene	0.094	0.08	0.00	0.07	3.568	-2.31	
1-octyne	0.155	0.22	0.09	0.10	3.521	-2.13	-4.49
chlorobenzene	0.718	0.65	0.00	0.07	3.657	-1.11	-4.02

Chapter 14 An algorithm for 323 Odour Detection Thresholds based on the Nagata data set

14.0. INTRODUCTION

Airborne substances can stimulate both the olfactory and the trigeminal nerve in the nose, giving rise to odour and pungent (irritant) sensations, respectively [1]. At the same time, these sensory responses are more amenable to quantification than other symptoms such as headache, tiredness or difficulty in concentration [2].

An odour detection threshold (ODT) is a biological endpoint that provides a quantitative assessment of the effect of airborne chemicals on the olfactory system of a human subject. Odour detection thresholds, obtained by the same protocol, of a series of chemicals then constitute a quantitative measure of the relative effectiveness, or potency, of the chemicals to elicit an effect. The smaller the ODT, the more 'potent' is the chemical. If odour detection thresholds are determined by two different protocols, it may be that, although relative potency remains the same, or roughly the same, for a series of chemicals, the absolute value of ODTs obtained by the two protocols may differ. Schmidt and Cain [3] have pointed out that in such a case, the protocol that gives rise to the lower absolute values might usually be regarded as the most acceptable. There are two reasons for this. First, a lower ODT could lead to a lower occupational exposure threshold, and from the point of view of risk assessment it is better to err on the side of the lower threshold. Second, it is more likely that any weakness in the methodology used to set up a protocol will result in higher thresholds than in lower thresholds.

There exists a growing set of data on the odour and nasal pungency thresholds of VOCs in humans. Substances tested include both individual members and mixtures of homologous series of alcohols, acetates, ketones, alkylbenzenes, and other miscellaneous compounds. In these series, physicochemical properties change in a systematic fashion. In principle, gathering data for such series with a uniform and standardized methodology would allow the construction of quantitative structure activity relationships (QSARs) for both odour and nasal pungency.

There are two recent protocols for the determination of ODTs that have both tested a relatively large number of chemicals ($N \geq 60$) and used a uniform methodology.

These uniform methodologies are the Japanese triangle odor bag method of Nagata [4], and the odour squeeze bottle method of Cometto-Muñiz and Cain [5,6]. Standard compilations of ODTs [7-9] have the drawback of retrieving values obtained with widely different methods and approaches. As has been pointed out [10], the determination of ODT values depend upon such factors as the method of stimulus dilution, volume of inhalation, type of psychophysical task, and number of trials presented, and requires the need for standardization of procedures. Both the procedures of Nagata and of Cometto-Muñiz and Cain involve standardized, explicit protocols.

Although the Nagata data set covers some 200 chemicals, and the Cometto-Muñiz and Cain data set includes 80 chemicals, the number of chemicals whose ODT values have been thus determined is but a small fraction of those known to be present in industrial and home environments. For example, the number of known compounds just in tobacco smoke exceeded 3,800 in 1982 [11]. It seems to be quite impractical to obtain experimental values of ODTs, at least by the current methodologies, for any other than a minority of known airborne pollutants.

The present work has two major aims. First to attempt to analyse the Nagata data set in order to provide an equation or algorithm that will enable the Nagata ODT values to be predicted, and will allow the estimation of ODT values for thousands of airborne chemicals. Second to examine the Nagata and Cometto-Muñiz and Cain data sets to see if they can be combined in any way, in order to extend the Nagata ODT values.

14.1. METHODOLOGY

A very general equation for the correlation of a variety of processes in which VOCs are transferred from the gas phase to some condensed phase has been devised by Abraham and co-workers which was dealt with in detail in Chapter 4 of this work.

$$SP = c + e \cdot \mathbf{E} + s \cdot \mathbf{S} + a \cdot \mathbf{A} + b \cdot \mathbf{B} + l \cdot \mathbf{L} \qquad (14.1)$$

In Eq. (14.1), the dependent variable, SP, is a set of solute properties in a given system. In the present case SP will be log (1/ODT) where ODT values are in ppm by volume. '1/ODT' is used instead of ODT so that the larger the value of log (1/ODT) the more potent is the chemical. The independent variables in Eq. (14.1) are as follows. **E** is the solute excess molar refraction, **S** is the dipolarity/polarizability, **A** and **B** are the overall

or effective hydrogen bond acidity and basicity and **L** is defined through L^{16}, the solute Ostwald solubility coefficient on hexadecane at 298K. The coefficients *c*, *e*, *s* , *a*, *b* and *l* reflect the complementary properties of the receptor phase. The *e*-coefficient gives the tendency of the phase to interact with VOCs through polarizability-type interactions, mostly via electron pairs. The *s*-coefficient is a measure of the phase dipolarity/polarizability. The *a*-coefficient represents the complementary property to VOC hydrogen-bond acidity and so is a measure of the phase hydrogen-bond basicity. Likewise, the *b*-coefficient is a measure of the phase hydrogen-bond acidity. Finally, the *l*-coefficient is a measure of the hydrophobicity of the phase.

14.2. RESULTS AND DISCUSSION

14.2.1 Correlation of the Nagata data set.

ODTs for a series of 223 compounds including esters, aldehydes, ketones, alcohols, carboxylic acids, aromatic hydrocarbons, terpenes, mercaptans, and a number of other VOCs, have been determined by Nagata [4], as mentioned, using the triangle odour bag method described in Chapter 2 of this work.

As a first step Eq. (14.1) was applied to all the VOCs for which the descriptors were already known, see Table 14.1, except 1-octene, 1-nonene, acetone, methyl formate, ethyl n-butyrate, ethyl isobutyrate, ethyl n-valerate, propane, butane, methanol, isopropanol and t-butanol which were revealed to be outliers in an initial study, see below section 14.2.1.1, and were removed to yield the correlation equation,

$$\text{Log (1/ODT)} = -1.71 + 1.29\,\mathbf{E} + 0.133\,\mathbf{S} + 0.429\,\mathbf{A} + 2.85\,\mathbf{B} + 0.512\,\mathbf{L} + 3.94\,\mathbf{M} + 1.92\,\mathbf{AL} + 1.64\,\mathbf{AC} + 1.18\,\mathbf{UE} \qquad (14.2)$$

$N = 193$, $R^2 = 0.718$, $SD = 0.886$, $F = 51.80$

The first problem that one can observe analyzing this data set is that the data are so scattered that it is not so easy to detect trends and especially to detect outliers. In addition to the usual descriptors, see Eq. (14.1), three additional descriptors were included that are 'indicator variables' to deal with specific classes that can be recognized as having a 'trend'; in all cases they are more potent than expected. **M** is the variable for the mercaptans and takes the value **M** = 1 for mercaptans and **M** = 0 for all

the other VOCs. **AL** is the variable for aldehydes and takes the value **AL** = 1 for aldehydes and **AL** = 0 for all the other VOCs. **AC** is the variable for acids and takes the value **AC** = 1 for acids and **AC** = 0 for all the other VOCs. **UE** is the variable for unsaturated esters and takes the value **UE** = 1 for unsaturated esters and **UE** = 0 for all other compounds.

Table 14.1
Nagata values of odour detection thresholds used to obtain Eq 14.2

Substance	Log (1/ODT)	Substance	Log (1/ODT)
sulfur dioxide	0.060	acetaldehyde	2.824
chlorine[a]	1.310	propionaldehyde[a]	3.000
n-pentane	-0.146	n-butylaldehyde	3.174
isopentane[a]	-0.114	isobutylaldehyde[a]	3.456
n-hexane	-0.176	n-valeraldehyde	3.387
2-methylpentane[a]	-0.845	isovaleraldehyde[a]	4.000
3-methylpentane	-0.949	n-hexylaldehyde	3.553
2,2-dimethylbutane[a]	-1.301	n-heptylaldehyde[a]	3.745
2,3-dimethylbutane	0.377	n-octylaldehyde	5.000
n-heptane[a]	0.174	n-nonylaldehyde[a]	3.469
2-methylhexane	0.377	n-decylaldehyde	3.398
3-methylhexane[a]	0.076	acrolein[a]	2.444
3-ethylpentane	0.432	crotonaldehyde	1.638
2,2-dimethylpentane[a]	-1.580	methacrolein[a]	2.071
2,3-dimethylpentane	-0.653	methyl ethyl ketone[a]	0.357
2,4-dimethylpentane[a]	0.027	methyl n-propyl ketone	1.553
n-octane	-0.230	methyl isopropyl ketone[a]	0.301
2-methylheptane[a]	0.959	methyl n-butyl ketone	1.620
3-methylheptane	-0.176	methyl sec.butyl ketone[a]	1.620
4-methylheptane[a]	-0.230	methyl isobutyl ketone	0.770
2,2,4-trimethylpentane	0.174	methyl tert.butyl ketone	1.367
n-nonane	-0.342	methyl n-amyl ketone	2.167
2,2,5-trimethylhexane	0.046	methyl isoamyl ketone[a]	2.678
n-decane[a]	0.208	diacetyl	4.301
n-undecane	0.060	ethyl formate	-0.431

Table 14.1 *(continued)*

Substance	Log (1/ODT)	Substance	Log (1/ODT)
n-dodecane[a]	0.959	n-propyl formate[a]	0.018
methylcyclopentane	-0.230	isopropyl formate	0.538
cyclohexane[a]	-0.398	n-butyl formate[a]	1.060
methylcyclohexane	0.824	isobutyl formate	0.310
propylene[a]	-1.114	methyl acetate[a]	-0.230
1-butene	0.444	ethyl acetate	0.060
isobutene[a]	-1.000	n-propyl acetate[a]	0.620
1-pentene	1.000	isopropyl acetate	0.796
1-hexene[a]	0.854	n-butyl acetate[a]	1.796
1-heptene	0.432	isobutyl acetate	2.097
1,3-butadiene[a]	0.638	sec.butyl acetate[a]	2.620
isoprene	1.319	tert.butyl acetate	1.149
chloroform[a]	-0.580	n-hexyl acetate[a]	2.745
carbon tetrachloride	-0.663	methyl propionate	1.009
trichloroethylene[a]	-0.591	ethyl propionate[a]	2.155
tetrachloroethylene	0.114	n-propyl propionate	1.237
formaldehyde[a]	0.301	isopropyl propionate[a]	2.387
n-butyl propionate	1.444	isobutanol[a]	1.959
isobutyl propionate[a]	1.699	sec.butanol	0.658
methyl n-butyrate	2.149	n-pentanol[a]	1.000
n-propyl n-butyrate[a]	1.959	sec.pentanol	0.538
isopropyl n-butyrate	2.208	isopentanol[a]	2.770
n-butyl n-butyrate[a]	2.319	tert.pentanol	1.056
isobutyl n-butyrate	2.796	n-hexanol[a]	2.222
methyl n-valerate[a]	2.658	n-heptanol	2.319
n-propyl n-valerate	2.481	n-octanol[a]	2.569
n-butyl isovalerate[a]	1.921	n-nonanol	3.046
methyl isobutyrate	2.721	n-decanol[a]	3.114
n-propyl isobutyrate[a]	2.699	2-ethoxyethanol	0.237
isopropyl isobutyrate	1.456	2-n-buthoxyethanol	1.367
n-butyl isobutyrate[a]	1.658	allyl sulfide	3.658

Table 14.1 *(continued)*

Substance	Log (1/ODT)	Substance	Log (1/ODT)
isobutyl isobutyrate	1.125	tetrahydrothiophene[a]	3.208
2-ethoxyethyl acetate[a]	1.310	carbon disulfide	0.678
acetonitrile	-1.114	benzene[a]	-0.431
acrylonitrile[a]	-0.944	toluene	0.481
ammonia	-0.176	ethylbenzene[a]	0.770
methylamine[a]	1.456	o-xylene	0.420
ethylamine	1.337	m-xylene[a]	1.387
n-propylamine	1.215	p-xylene	1.237
isopropylamine	1.602	n-propylbenzene[a]	2.420
n-butylamine[a]	0.770	isopropylbenzene	2.076
isobutylamine	2.824	1,2,4-trimethylbenzene[a]	0.921
sec.butylamine[a]	0.770	1,3,5-trimethylbenzene	0.770
tert.butylamine	0.770	o-ethyltoluene[a]	1.131
dimethylamine[a]	1.481	m-ethyltoluene	1.745
diethylamine	1.319	p-ethyltoluene[a]	2.081
trimethylamine	4.495	n-butylbenzene	2.071
triethylamine[a]	2.268	o-diethylbenzene[a]	2.027
acetic acid[a]	2.222	m-diethylbenzene	1.155
propionic acid	2.244	p-diethylbenzene[a]	3.409
n-butyric acid[a]	3.721	1,2,3,4-tetramethylbenzene	1.959
isobutyric acid	2.824	styrene[a]	1.456
n-valeric acid[a]	4.432	phenol	2.252
isovaleric acid	4.108	o-cresol[a]	3.553
n-hexanoic acid[a]	3.222	m-cresol	4.000
ethanol	0.284	p-cresol[a]	4.268
n-propanol[a]	1.027	furane	-0.996
n-butanol	1.420	pyridine	1.201
sec.butyl mercaptane[a]	4.523	ß-pinene[a]	1.481
hydrogen sulfide	3.387	limonene	1.420
methyl mercaptane[a]	4.155	n-butyl acrylate[a]	3.260
ethyl mercaptane	5.060	isobutyl acrylate	3.046

Table 14.1 *(continued)*

Substance	Log (1/ODT)	Substance	Log (1/ODT)
n-propyl mercaptane[a]	4.886	tert.butyl mercaptane[a]	4.538
isopropyl mercaptane	5.222	n-amyl mercaptane	6.108
n-butyl mercaptane[a]	5.553	isoamyl mercaptane[a]	6.114
isobutyl mercaptane	5.167	n-hexyl mercaptane	4.824
methyl acrylate[a]	2.456	dimethyl sulfide[a]	2.523
ethyl acrylate	3.585	diethyl sulfide	4.481
indole	3.523	dimethyl disulfide[a]	2.658
skatole[a]	5.252	diethyl disulfide	2.699
thiophene	3.252		

[a]Part of the training set used for Eq. 14.9; see also Tables 14.7 and 14.8.

14.2.1.1. Initial analysis

The statistics of the final equation with N= 193, $R^2 = 0.705$, R = 0.840, SD = 0.903 and F = 55.06 seem not to be all that good with an SD nearly one log unit. However a number of charts are given to illustrate the scatter of the data, as follows. For any homologous series, all the descriptors in Eq. (14.2) are approximately constant except for **L**. Then Eq. (14.2) reverts to the simple equation

$$\log(1/ODT) = c' + 0.512\,\mathbf{L} \qquad (14.3)$$

where c' will include all the other terms and is constant for a homologous series. Then for each homologous series a plot of log (1/ODT) against **L** will have a known slope of 0.512 and a given intercept of c'. If the data are very 'good' the plot will have all the points for a given series on the calculated line but if the data are 'bad' the points will be very scattered about the line, see Figure 14.1.

The first such plot (Figure 14.1a) is for the alkanes. It can be seen that there is considerable scatter about the calculated line, showing that the data are not good at all, but also that there are two alkanes, butane and propane, shown by circles, that are so far out that they are considered as outliers and have been left out of Eq. (14.2). The second plot is for alkenes (Figure 14.1b). Again there is a large scatter and two points, 1-octene and 1-nonene, were considered to be outliers. The third plot is for esters (Figure 14.1c).

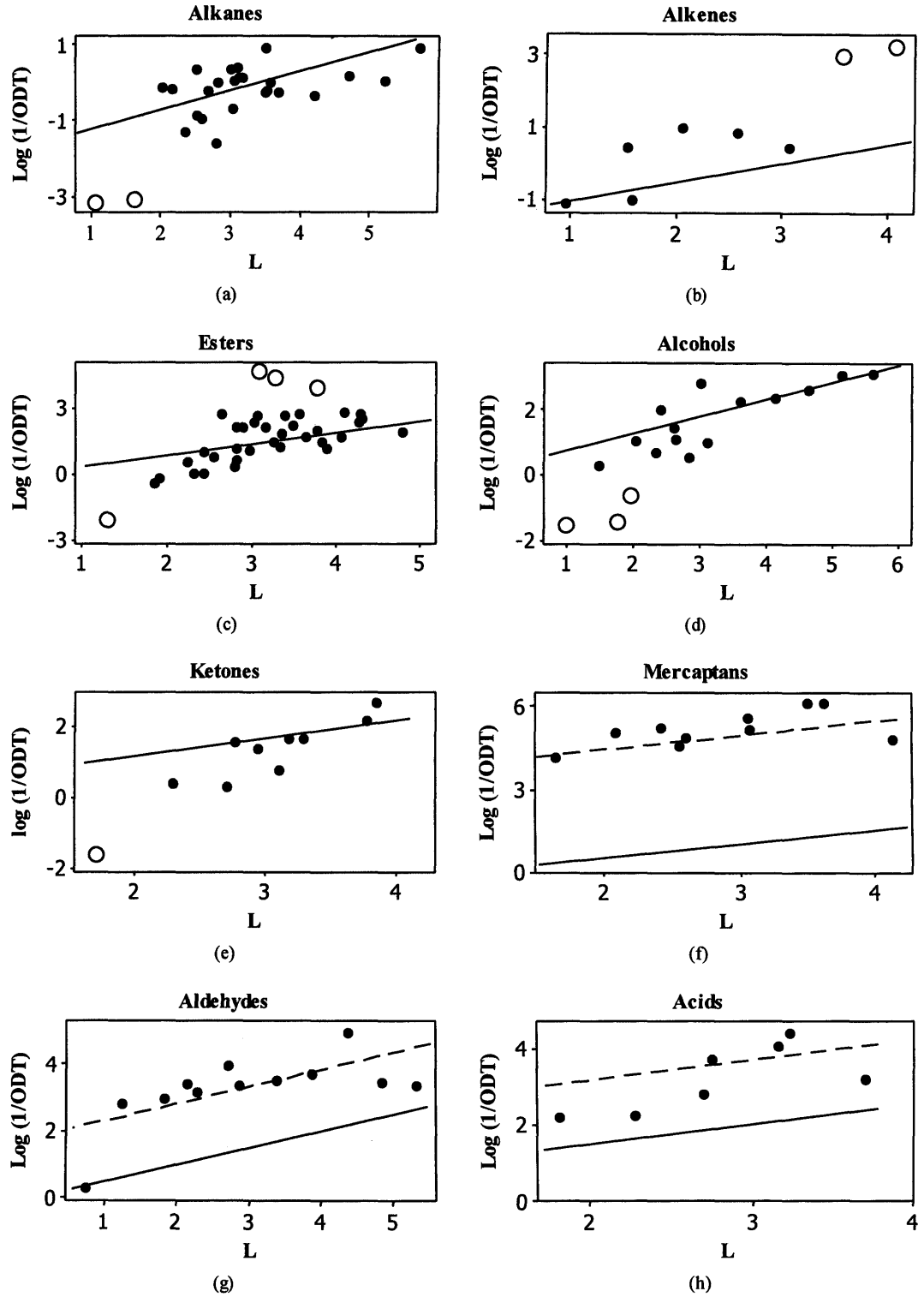


Figure 14.1. Scatterplot of log (1/ODT) against **L** for eight different homologous series of VOCs. Calculated line without the indicator variable (———). Calculated line with the indicator variable (– – – –). Outliers (○).

This seems quite good, but if the different Y-scale is observed, then it can be seen that there is a large scatter once again. Here there are four points that have been taken as outliers, namely methyl formate, ethyl n-butyrate, ethyl isobutyrate, and ethyl

n-valerate. The fourth plot is for alcohols (Figure 14.1d), and again as shown by the circles there are a number of outliers: methanol, isopropanol, and tert-butanol. The fifth plot for ketones (Figure 14.1e) has one outlier, acetone. The sixth plot for the mercaptans (Figure 14.1f) is very interesting, because this clearly shows that they are all much more potent than calculated. This is a clear trend and is not the usual scatter. On this plot is given the calculated line without the indicator variable, **M**, shown with a continuous line, and the calculated line with the indicator variable shown with a dashed line. Now it is very clear that an indicator variable is needed, which can bring the mercaptans into line. The seventh (Figure 14.1g) and eighth (Figure 14.1h) plots are for the aldehydes and carboxylic acids, and show the same effect as for the mercaptans, namely that the indicator variable brings them into line. The results for the aldehydes and carboxylic acids are exactly as we found previously for the Abraham and co-workers data set [6]. Finally, unsaturated esters are another class of compounds that are more potent than calculated from Eq. (14.2) (graph not shown).

The analysis of the residuals of these 193 compounds shows a couple of points, i.e. diacetyl and trimethylamine, very far from the rest. These two compounds were then considered to be outliers and also left out, leading to the very final equation,

$$\text{Log (1/ODT)} = -1.83 + 0.882\,\mathbf{E} + 0.408\,\mathbf{S} + 0.999\,\mathbf{A} + 2.200\,\mathbf{B} + 0.578\,\mathbf{L} + 4.06\,\mathbf{M} + 1.80\,\mathbf{AL} + 1.42\,\mathbf{AC} + 1.29\,\mathbf{UE} \quad (14.4)$$

$N = 191$, $R^2 = 0.748$, SD = 0.830, F = 59.75

In the analysis of odour thresholds by Abraham and co-worker, already described in Chapter 12, a parabolic term in $\mathbf{L}^2$ was included as well as the indicator variables, **AC** and **AL,** for acids and aldehydes, respectively. The same parameter was included in Eq. (14.5) with the aim of comparing with Eq. (12.5)

$$\text{Log (1/ODT)} = -2.01 + 0.895\,\mathbf{E} + 0.418\,\mathbf{S} + 1.065\,\mathbf{A} + 2.165\,\mathbf{B} + 0.707\,\mathbf{L} - 0.020\,\mathbf{L}^2 + 4.062\,\mathbf{M} + 1.826\,\mathbf{AL} + 1.379\,\mathbf{AC} + 1.282\,\mathbf{UE} \quad (14.5)$$

$N = 191$, $R^2 = 0.749$, SD = 0.834, F = 53.57

The statistics of Eq. (14.5) show however that the extra parameter $\mathbf{L}^2$ is not significant, with a P value of 0.632 and that its introduction in the equation does not improve the correlation at all. This suggests that the model without that coefficient may be more appropriate.

The coefficients in Eq. (14.4) can be compared with those for various gas-condensed phase partitions that take place by simple transfer mechanisms, see Table 14.2 [12-15]. There is considerable similarity between Eq. (14.4) and equations for the solubility of gaseous VOCs in solvents such as wet 1-octanol and dry methanol. There is also some similarity with equations for the solubility of gaseous VOCs in a number of biophases [16].

Table 14.2
Regression coefficients in Eq. (14.1) for gas-solvent phase partitions at 298K

Phase	e	s	a	b	l
wet 1-octanol	0.002	0.709	3.519	1.429	0.858
dry methanol	-0.215	1.173	3.701	1.432	0.769
chloroform	-0.467	1.203	0.138	1.432	0.994
acetone	-0.277	1.522	3.258	0.078	0.863
dimethylformamide	-0.189	2.327	4.756	0.000	0.808
water	0.822	2.743	3.904	4.814	-0.213
brain[a]	0.427	0.286	2.781	2.787	0.609
muscle[a]	0.544	0.216	3.471	2.924	0.578
fat[a]	-0.172	0.729	1.747	0.219	0.895
ODT[a]	0.533	1.912	1.276	1.559	0.699
ODT, Eq. (14.4)	0.882	0.408	0.999	2.200	0.578

[a]At 310K. Adapted from ref. [6]

Therefore, the statistics of Eq (14.4) can be regarded as reasonable, especially since the self-consistency of Nagata's data is around 0.66 log unit. In Table 14.3 are shown those compounds for which the consistency has been studied. The study includes 80 compounds, i.e. alcohols, alkanes, esters and ketones. The average error, AE, between the observed (Olog (1/ODT)) and calculated (Clog (1/ODT)) values is -0.06 log units, the average absolute error, AAE, is 0.54 log units and the standard desviation, S, is 0.66 log units. The very small AE shows that there is little bias in the assignments of Eq. (14.4), but the large values of AAE and S indicate that there is a considerable inconsistency in the Nagata data. One cannot claim that the AAE and S values are a result of experimental errors, because they are derived from the fitting of Eq. (14.1) to Nagata's data. What one can deduce is that any equation constructed to correlate

Nagata's data will not have an AAE value less than about 0.54 log units or an SD value of less than 0.66 log units, unless the equation is seriously over fitted.

Table 14.3
Compounds used to study the self-consistency of the Nagata data set

Substance	Obs log (1/ODT)	Calc log (1/ODT)	Obs-Calc	|Obs-Calc|	(Obs-Calc)^2
ethanol	0.284	0.977	-0.693	0.693	0.480
n-propanol	1.027	1.256	-0.229	0.229	0.052
n-butanol	1.420	1.547	-0.127	0.127	0.016
isobutanol	1.959	1.451	0.508	0.508	0.258
sec.butanol	0.658	1.413	-0.755	0.755	0.570
n-pentanol	1.000	1.805	-0.805	0.805	0.648
sec.pentanol	0.538	1.669	-1.132	1.132	1.281
isopentanol	2.770	1.757	1.013	1.013	1.026
tert.pentanol	1.056	1.562	-0.506	0.506	0.256
n-hexanol	2.222	2.063	0.159	0.159	0.025
n-heptanol	2.319	2.321	-0.002	0.002	0.000
n-octanol	2.569	2.578	-0.010	0.010	0.000
n-nonanol	3.046	2.834	0.211	0.211	0.045
n-decanol	3.114	3.084	0.029	0.029	0.001
n-pentane	-0.15	-0.595	0.445	0.445	0.198
isopentane	-0.11	-0.671	0.561	0.561	0.315
n-hexane	-0.18	-0.337	0.157	0.157	0.025
2-methylpentane	-0.85	-0.421	-0.429	0.429	0.184
3-methylpentane	-0.95	-0.381	-0.569	0.569	0.324
2,2-dimethylbutane	-1.3	-0.498	-0.802	0.802	0.643
2,3-dimethylbutane	0.38	-0.425	0.805	0.805	0.648
n-heptane	0.17	-0.079	0.249	0.249	0.062
2-methylhexane	0.38	-0.166	0.546	0.546	0.299
3-methylhexane	0.08	-0.145	0.225	0.225	0.050
3-ethylpentane	0.43	-0.120	0.550	0.550	0.303
2,2-dimethylpentane	-1.58	-0.271	-1.309	1.309	1.713
2,3-dimethylpentane	-0.65	-0.159	-0.491	0.491	0.241

Table 14.3 *(continued)*

Substance	Obs log (1/ODT)	Calc log (1/ODT)	Obs-Calc	|Obs-Calc|	(Obs-Calc)^2
2,4-dimethylpentane	0.03	-0.265	0.295	0.295	0.087
n-octane	-0.23	0.179	-0.409	0.409	0.167
2-methylheptane	0.96	0.078	0.882	0.882	0.777
3-methylheptane	-0.18	0.094	-0.274	0.274	0.075
4-methylheptane	-0.23	0.080	-0.310	0.310	0.096
2,2,4-trimethylpentane	0.17	-0.113	0.283	0.283	0.080
n-nonane[a]	-0.34	0.437	-0.777	0.777	0.604
2,2,5-trimethylhexane	0.05	0.123	-0.073	0.073	0.005
n-decane	0.21	0.695	-0.485	0.485	0.235
n-undecane	0.06	0.953	-0.893	0.893	0.797
n-dodecane	0.96	1.211	-0.251	0.251	0.063
ethyl formate	-0.431	0.783	-1.214	1.214	1.474
n-propyl formate	0.018	1.083	-1.066	1.066	1.135
isopropyl formate	0.538	0.980	-0.442	0.442	0.195
n-butyl formate	1.060	1.352	-0.291	0.291	0.085
isobutyl formate	0.310	1.265	-0.955	0.955	0.913
methyl acetate	-0.230	0.817	-1.047	1.047	1.096
ethyl acetate	0.060	1.022	-0.962	0.962	0.925
n-propyl acetate	0.620	1.281	-0.661	0.661	0.437
isopropyl acetate	0.796	1.141	-0.345	0.345	0.119
n-butyl acetate	1.796	1.553	0.242	0.242	0.059
isobutyl acetate	2.097	1.455	0.642	0.642	0.412
sec.butyl acetate	2.620	1.401	1.219	1.219	1.486
tert.butyl acetate	1.149	1.272	-0.123	0.123	0.015
n-hexyl acetate	2.745	2.032	0.713	0.713	0.508
methyl propionate	1.009	1.082	-0.073	0.073	0.005
ethyl propionate	2.155	1.274	0.881	0.881	0.775
n-propyl propionate	1.237	1.546	-0.309	0.309	0.096
isopropyl propionate	2.387	1.387	1.000	1.000	1.000
n-butyl propionate	1.444	1.799	-0.355	0.355	0.126
isobutyl propionate	1.699	1.697	0.001	0.001	0.000

Table 14.3 *(continued)*

Substance	Obs log (1/ODT)	Calc log (1/ODT)	Obs-Calc	|Obs-Calc|	(Obs-Calc)^2
methyl n-butyrate	2.149	1.318	0.830	0.830	0.690
n-propy n-butyrate	1.959	1.773	0.185	0.185	0.034
isopropyl n-butyrate	2.208	1.619	0.588	0.588	0.346
n-butyl n-butyrate	2.319	2.025	0.294	0.294	0.087
isobutyl n-butyrate	2.796	1.934	0.862	0.862	0.744
methyl n-valerate	2.658	1.573	1.084	1.084	1.176
n-propyl n-valerate	2.481	2.043	0.438	0.438	0.192
n-butyl isovalerate	1.921	2.290	-0.369	0.369	0.136
methyl isobutyrate	2.721	1.187	1.534	1.534	2.354
n-propyl isobutyrate	2.699	1.657	1.042	1.042	1.087
isopropyl isobutyrate	1.456	1.500	-0.044	0.044	0.002
n-butyl isobutyrate	1.658	1.919	-0.261	0.261	0.068
isobutyl isobutyrate	1.125	1.825	-0.700	0.700	0.490
methyl ethyl ketone	0.357	1.270	-0.913	0.913	0.834
methyl n-propyl ketone	1.553	1.509	0.044	0.044	0.002
methyl isopropyl ketone	0.301	1.477	-1.176	1.176	1.382
methyl n-butyl ketone	1.620	1.780	-0.160	0.160	0.026
methyl sec.butyl ketone	1.620	1.717	-0.098	0.098	0.010
methyl isobutyl ketone	0.770	1.679	-0.910	0.910	0.828
methyl tert.butyl ketone	1.367	1.597	-0.231	0.231	0.053
methyl n-amyl ketone	2.167	2.022	0.145	0.145	0.021
methyl isoamyl ketone	2.678	2.061	0.617	0.617	0.381
			AE	**AAE**	**SD**
			-0.06	**0.54**	**0.66**

The indicator variables used in Eq. (14.4) are not just arbitrary variables used to obtain a better fit to the data; similar excess potency has been described before. Alarie et al. [17,18] investigated the sensory irritation of mice by airborne chemicals, and classed chemicals as acting by a physical mechanism (*p*) or by a chemical mechanism (*c*). Compounds that induced sensory irritation by a chemical mechanism were identified through an increase in potency by comparison with that calculated for irritation by a

physical mechanism. In essence, this is the same procedure that we have used to identify compounds that are more potent than calculated from Eq. (14.3). Alarie et al. [17,18] showed that carboxylic acids, aldehydes, and unsaturated esters were more potent than expected, exactly as we have done. Mercaptans were not studied by Alarie et al. [17,18] but other potent compounds, not studied by Nagata [4], were unsaturated ketones, and isocyanates. The only compounds for which our findings and those of Alarie et al. are different are the amines which Alarie et al. found to be more potent than expected, whereas we find that amines fit Eq. (14.3) quite well.

14.2.2 Comparison of the Nagata and Cometto-Muñiz and co-workers data set.

In order to compare the data set by Nagata [4] with the one by Cometto-Muñiz and co-workers [6] it is instructive to plot the odour threshold values of 39 common compounds found in both data sets (Figure 14.2), summarized in Table 14.4. From Figure 14.2, it may be concluded that the two sets of data are not very comparable. There is a large scatter and the regression line does not go through the origin.

Table 14.4
Common compounds in the Nagata and Cometto-Muñiz and co-workers data set

	log (1/ODT)			log (1/ODT)	
Stimuli	**Nagata**	**Cometto-Muñiz**	**Stimuli**	**Nagata**	**Cometto-Muñiz**
1-octene	3.000	-2.310	ethanol	0.284	-1.845
n-butylaldehyde	3.174	-0.477	n-propanol	1.027	-1.146
n-valeraldehyde	3.387	-0.699	isopropanol	-1.415	-2.702
n-hexylaldehyde	3.553	1.097	n-butanol	1.420	-0.301
n-heptylaldehyde	3.745	1.523	sec.butanol	0.658	-1.978
n-octylaldehyde	5.000	2.398	tert.butanol	-0.653	-2.785
acetone	-1.623	-4.065	n-pentanol	1.000	-0.114
methyl n-propyl ketone	1.553	-0.935	n-hexanol	2.222	0.046
methyl n-amyl ketone	2.167	0.149	n-heptanol	2.319	1.000
methyl acetate	-0.230	-3.459	n-octanol	2.569	2.155
ethyl acetate	0.060	-2.238	toluene	0.481	-2.190
n-propyl acetate	0.620	-1.394	ethylbenzene	0.770	-1.260

Table 14.4 *(continued)*

Stimuli	log (1/ODT) Nagata	log (1/ODT) Cometto-Muñiz	Stimuli	log (1/ODT) Nagata	log (1/ODT) Cometto-Muñiz
n-butyl acetate	1.796	-0.378	n-propylbenzene	2.420	-0.467
sec.butyl acetate	2.620	-0.672	isopropylbenzene	2.076	-0.033
tert.butyl acetate	1.149	-0.114	n-butylbenzene	2.071	-0.629
n-hexyl acetate	2.745	0.201	pyridine	1.201	-0.107
acetic acid	2.222	2.000	a-pinene	1.745	-1.277
n-butyric acid	3.721	2.444	ß-pinene	1.481	-1.070
n-hexanoic acid	3.222	2.585	limonene	1.420	-0.659
methanol	-1.519	-3.176			

The differences in the observed values of log (1/ODT) by Nagata against the observed values of log (1/ODT) by Cometto-Muñiz and co-workers are shown in Figure 14.3. Carboxylic acids and aliphatic aldehydes were left out because of the indicator variable needed for these compounds, leaving us with 31 common compounds. In the first comparison an outlier was found, i.e. 1-octene (shown surrounded by a circle in Figure 14.3), and left out to carry out a better analysis.

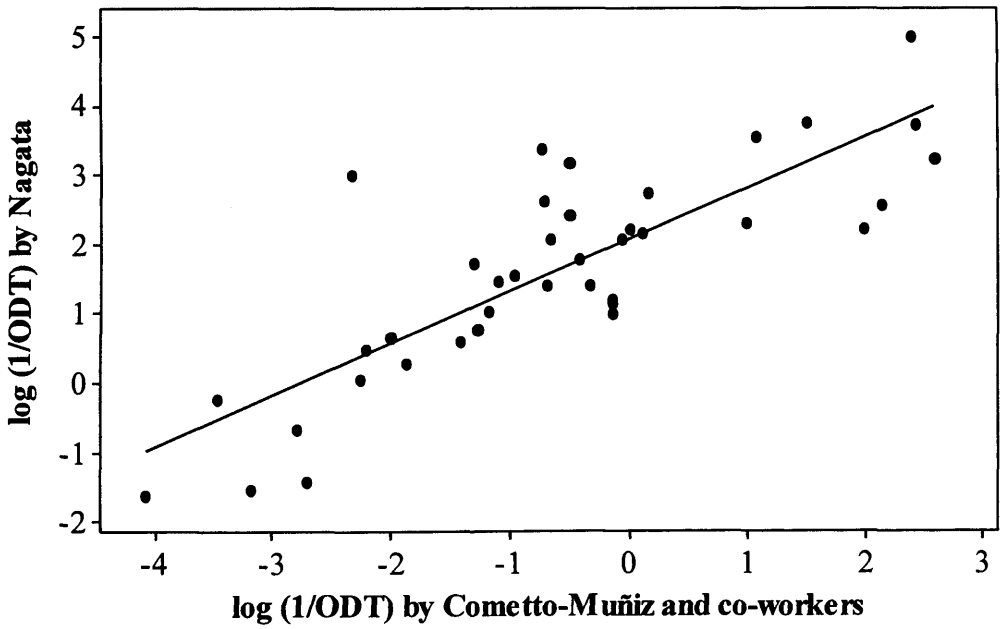


Figure 14.2. Plot of observed values of log (1/ODT) by Nagata against the observed values of log (1/ODT) by Cometto-Muñiz and co-workers.

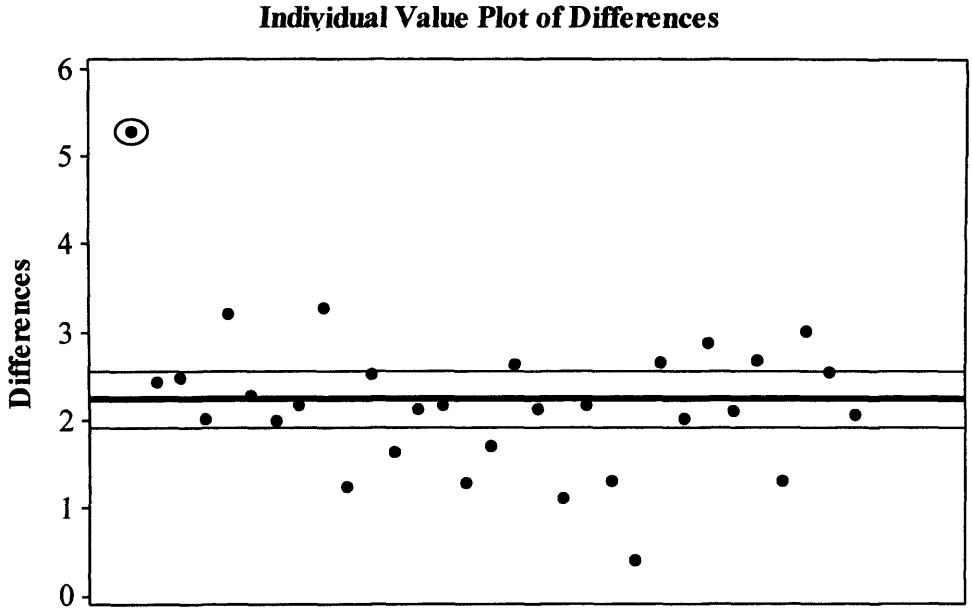


Figure 14.3. Difference in the observed values of log (1/ODT) by Nagata and the observed values of log (1/ODT) by Cometto-Muñiz and co-workers. The average value of the differences and the confidence interval at 95% are included.

As a result of the second analysis, see Figure 14.4, the average value of the differences in log (1/ODT) is 2.129 between the Nagata and the Cometto-Muñiz and co-workers data set, and the confidence interval at 95% is (1.883, 2.375). Statistically when the

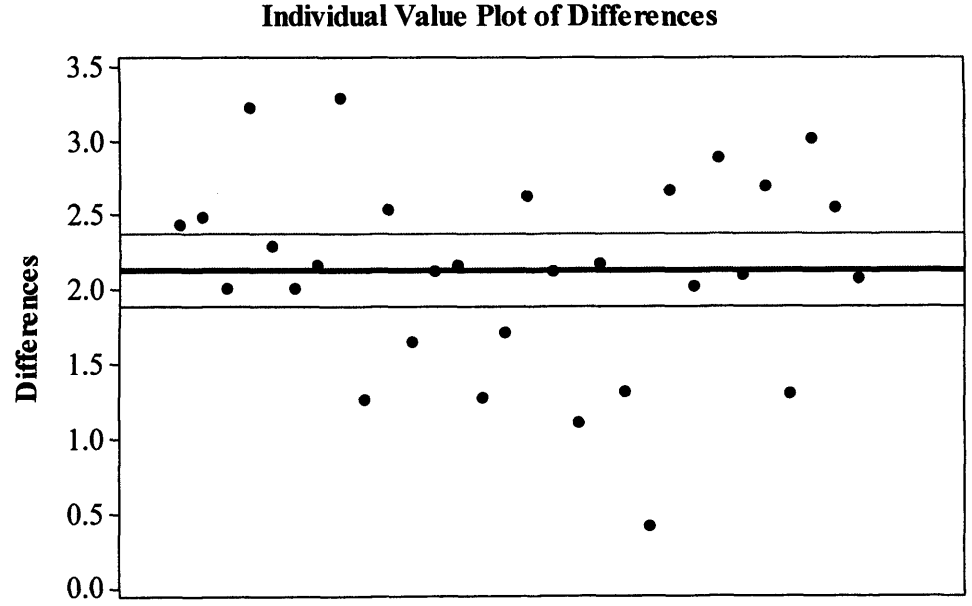


Figure 14.4. Difference in the observed values of log (1/ODT) by Nagata and the observed values of log (1/ODT) by Cometto-Muñiz and co-workers after leaving out the outlier, 1-octene. The average value of the differences and the confidence interval at 95% are included.

confidence interval does not include the value zero, it means that the average of both data sets is not the same. In other words, the results of the two groups are not comparable. This can be due to the protocol followed at the time of doing their experiments, which were quite different as described in Chapter 2 of this work. In Figure 14.4, it can be observed that the differences are all shifted in the same direction. The average value of the differences, 2.129, can be subtracted from the log (1/ODT) values by Nagata and compared again with those by Cometto-Muñiz and co-workers.

As expected, see Figure 14.5, the average of the differences is now 0.000 and the confidence interval at 95% is (-0.246, 0.246), which includes the value zero. This means that the average of log (1/ODT) values for both data sets is now the same. Therefore, it can be concluded that for the 30 compounds tested by this Paired t-test there is an average systematic difference of 2.129 log units between the Nagata and the Cometto-Muñiz and co-workers data set. If the absolute value of differences of the latest Paired t-test is averaged, a mean random error of 0.501 log units is found between the data sets. This is an important quantity because it implies that any correlation equation that includes log (1/ODT) values from both of the data sets cannot have a regression SD of less that 0.50 log units. This mean random error could be in either the Nagata data set or the Cometto-Muñiz and co-workers data set.

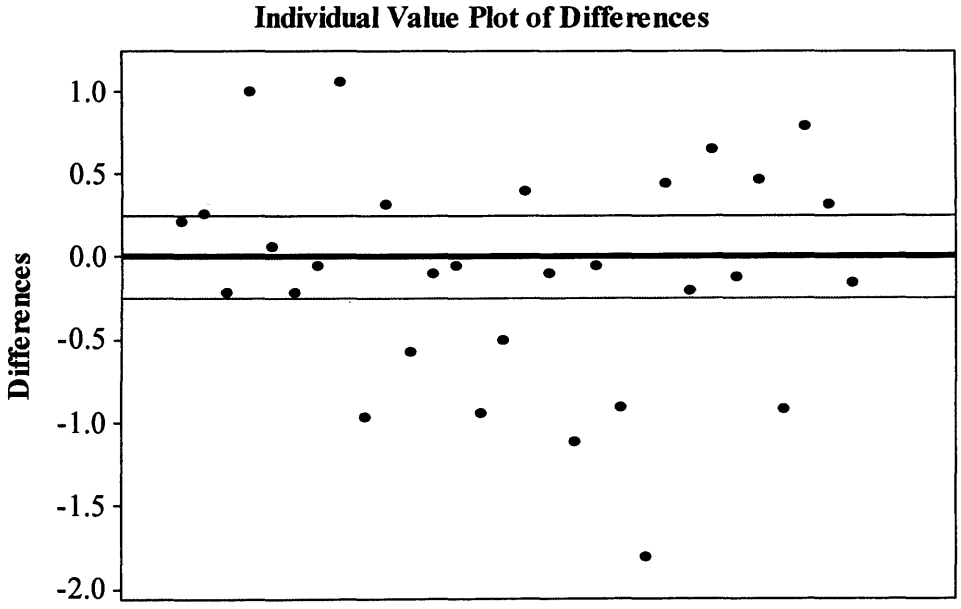


Figure 14.5. Difference in the observed values of log (1/ODT) by Nagata and the observed values of log (1/ODT) by Cometto-Muñiz and co-workers after leaving out the outlier, 1-octene, and subtracting the value 2.129 of the log (1/ODT) values by Nagata. The average value of the differences and the confidence interval at 95% are included.

These results make it necessary to investigate whether there is a systematic error in the measurements or just random error for the data set of each group. However, in the absence of individual values for the data set by Nagata [10], it is only possible to carry out an error analysis for the data set by Cometto-Muñiz and Cain [5], see Table 14.5. This analysis was carried out using the same number of compounds, 60, that was used by Cometto-Muñiz and co-workers [6], in their analysis of the data of Cometto-Muñiz.

Table 14.5
Four replicate values of log (1/ODT) with ODT in p.p.m

Stimuli	log 1/ODT (1)	log 1/ODT (2)	log 1/ODT (3)	log 1/ODT (4)
methanol	-2.95	-3.4	-2.6	-3.88
ethanol	-1.46	-2.63	-1.18	-2.36
1-propanol	-0.67	-1.54	-0.54	-1.43
2-propanol	-2.72	-3.28	-2.72	-2.08
1-butanol	-0.38	-0.18	0.022	-0.48
2-butanol	-1.93	-1.85	-2.68	-1.45
2-methyl-2-propanol	-2.95	-2.23	-3.72	-2.23
1-pentanol	-0.18	-0.079	0.4	-0.78
1-hexanol	0.26	0.2	0.4	-0.76
1-heptanol	1.12	0.64	1.82	0.22
4-heptanol	-0.88	-0.88	-1.3	-0.6
1-octanol	2	1.82	3.6	1.15
methyl acetate	-2.98	-3.3	-2.95	-4.6
ethyl acetate	-2.3	-2.3	-1.45	-2.9
propyl acetate	-1.48	-1.48	-0.78	-1.85
butyl acetate	-0.26	-0.78	0.3	-0.78
sec-butyl acetate	-1.23	-1.3	-1.68	-1.11
tert-butyl acetate	-0.73	-0.36	-2.36	0.43
pentyl acetate	0.046	-0.38	0.22	-0.18
hexyl acetate	0.3	-0.18	0.85	-0.18
heptyl acetate	0.046	0.046	0.49	-0.53
octyl acetate	0.66	0.12	1.15	-0.3
decyl acetate	1	-0.041	1.43	-0.4

Table 14.5 *(continued)*

Stimuli	log 1/ODT (1)	log 1/ODT (2)	log 1/ODT (3)	log 1/ODT (4)
dodecyl acetate	2.19	0.85	2.35	0.022
2-propanone	-3.43	-4.6	-4.11	-4.11
2-pentanone	-0.98	-0.98	-1.3	-0.48
2-heptanone	-0.3	-0.079	-0.6	-0.079
2-nonanone	-0.26	0.23	-0.6	0.74
toluene	-2.23	-1.58	-2.72	-2.23
ethyl benzene	-0.52	-0.82	-2.45	-1.26
propyl benzene	-0.26	0.43	-1.64	-0.4
butyl benzene	-0.41	0.2	-1.38	-0.92
pentyl benzene	-0.097	0.74	-1	0.37
hexyl benzene	0.54	0.54	-0.4	0.097
heptyl benzene	-0.11	1.05	0.19	-0.11
octyl benzene	1.05	0.68	-0.65	0.68
butanal	0.0044	-2.52	-0.72	0.96
pentanal	-0.74	-0.84	-1.81	-0.53
hexanal	0.68	1.19	0.68	2.8
heptanal	1.74	1.44	1.66	1.44
octanal	1.82	1.4	1.82	3.59
formic acid	-1.66	-0.11	-1.61	-1.92
acetic acid	1.42		1.92	3
butanoic acid	1.64	1.89	2.12	2.89
hexanoic acid	2.09	2.68	2.72	3.59
octanoic acid	4.82	3.96	3.8	6.6
cumene	-0.56	-0.12	0.33	0.22
p-cymene	-0.2	-0.66	-0.49	0.88
Δ-3-Carene	-0.41	0.06	-0.3	-0.24
linalool	-0.47	-0.57	0.02	1.11
1,8-Cineole	0.85	0.56	0.26	0.32
geraniol	0.71	0.26	1.69	1.6
α-terpinene	0.5	0.17	-0.69	-0.58

Table 14.5 *(continued)*

Stimuli	log 1/ODT (1)	log 1/ODT (2)	log 1/ODT (3)	log 1/ODT (4)
γ-terpinene	-0.26	-1.2	-1.35	-1.15
α-pinene	-1.14	-1.24	-1.44	-1.29
β-pinene	-0.83	-0.67	-1.63	-1.15
(R)-(+)- limonene	-0.9	-0.95	-1.22	-0.9
(S)- (-)-limonene	-0.04	-0.72	-1.35	-0.51
β-phenyl ethyl alcohol	1.74	2.3	3.46	1.26
pyridine	-0.15	0	0	-0.28
menthol	1.19	1.7	2.59	1.19
1-octene	-1.75	-1.53	-4	-1.97
1-octyne	-1.7	-1.26	-3.64	-1.93
chlorobenzene	-1.11	-0.6	-1.86	-0.87

As a result of the analysis of the four replicate values for each compound in the data set by Cometto-Muñiz and Cain [5], see Table 14.6, it can be concluded that the error in the regression line is principally due to a lack of fit of the regression equation and not due to experimental errors. The key figures in Table 14.6 are MS (Mean Square) for lack of fit (1.435) and MS for pure error (0.433), which indicate that the lack of fit error is much larger than the pure error, the actual error in the data. This is quantified by F = 3.31 which is the ratio between them (1.435/0.433). This means that the chance of the pure error being more than the error in the lack of fit is less than 5% or, the other way round, that the chance of the lack of fit error being more than the pure error is larger than 95%.

Table 14.6
Analysis of Variance for the replicates of the data by Cometto-Muñiz and Cain [5]

Source	DF	SS	MS	F	P
Regression	7	422.466	60.352	92.44	0.000
Residual Error	228	148.851	0.653		
Lack of Fit	50	71.729	1.435	3.31	0.000
Pure Error	178	77.122	0.433		
Total	235	571.317			

The serial procedure is to combine the two data sets to see how the statistics of the equation are. The Nagata absolute ODT values are lower than the Cometto-Muñiz and co-workers values by a factor of about 100. Following the suggestion of Schmidt and Cain [3], it would seem appropriate to set out a combined scale based on the Nagata data with the smaller threshold values. The equation obtained and some of its statistics are indicated below. A new indicator variable, **C**, was introduced to overcome the difference between the common compounds in the two data sets, and place Cometto-Muñiz and co-workers data on the Nagata scale. **C** takes the value 1 for the Cometto-Muñiz and co-workers log (1/ODT) values and 0 for the Nagata log (1/ODT) values. Aldehydes and acids, from Cometto-Muñiz and co-workers' data set, need again an indicator variable, but in this case one for each one. For acids, **CAC** was used and for aldehydes, **CAL**.

Table 14.7
Cometto-Muñiz and co-workers values of odour detection thresholds used in Eq. 14.6

Substance	Log (1/ODT)	Substance	Log (1/ODT)
1-octene	-1.750	2-methyl-2-propanol[a]	-2.783
1-octyne[a]	-1.630	1-pentanol	-0.160
butanal[a]	0.081	1-hexanol[a]	0.025
pentanal	-0.703	1-heptanol	0.950
hexanal[a]	0.850	4-heptanol[a]	-0.915
heptanal	1.570	toluene[a]	-2.190
octanal[a]	1.680	ethyl benzene	-1.263
2-pentanone	-0.935	propyl benzene[a]	-0.468
2-heptanone[a]	-0.265	cumene[a]	-0.033
2-nonanone	0.028	butyl benzene	-0.628
ethyl acetate	-2.238	p-cymene	-0.118
propyl acetate[a]	-1.398	pentyl benzene[a]	0.003
butyl acetate	-0.380	hexyl benzene	0.194
sec-butyl acetate[a]	-1.330	heptyl benzene[a]	0.255
pentyl acetate	-0.074	octyl benzene	0.440
hexyl acetate[a]	0.198	chlorobenzene	-1.110
heptyl acetate	0.013	β-phenyl ethyl alcohol[a]	2.190
octyl acetate[a]	0.408	pyridine	-0.108

Table 14.7 *(Continued)*

Substance	Log (1/ODT)	Substance	Log (1/ODT)
decyl acetate	0.497	α-pinene[a]	-1.278
dodecyl acetate[a]	1.353	β-pinene	-1.070
formic acid[a]	-1.730	(R)-(+)- limonene[a]	-0.993
butanoic acid[a]	2.135	(S)- (-)-limonene	-0.655
hexanoic acid	2.770	α –terpinene[a]	-0.150
octanoic acid[a]	4.193	γ-terpinene	-0.990
methanol[a]	-3.208	Δ-3-carene[a]	-0.223
ethanol	-1.908	1,8-cineole[a]	0.498
1-propanol[a]	-1.045	linalool	-0.340
2-propanol	-2.700	geraniol	1.333
1-butanol[a]	-0.255	menthol[a]	1.668
2-butanol	-1.978		

[a]Part of the training set used for Eq. (14.9); see also Tables 14.1 and 14.8.

As can be seen, the value of the coefficient for the indicator variable **C** (-2.471) is quite close to the average difference of 2.129 log units between the two data sets for the 30 compounds. The coefficient **C** is negative because what is correlated is the log (1/ODT). The small difference is, of course, because in this case, this indicator variable is not alone but in the middle of the equation and it is affected by the rest of the coefficients. The final equation that includes the Nagata data and the Cometto-Muñiz and co-workers data is,

$$\begin{aligned}\text{Log (1/ODT)} = &- 2.066 + 0.624\,\mathbf{E} + 0.623\,\mathbf{S} + 0.922\,\mathbf{A} + 2.127\,\mathbf{B} + 0.649\,\mathbf{L} \\ &+ 4.139\,\mathbf{M} + 1.789\,\mathbf{AL} + 1.461\,\mathbf{AC} + 1.264\,\mathbf{ES} - 2.471\,\mathbf{C} \\ &+ 3.183\,\mathbf{CAC} + 1.675\,\mathbf{CAL}\end{aligned} \qquad (14.6)$$

$N = 250$, $R^2 = 0.817$, $SD = 0.775$, $F = 87.94$

To reach to this Eq. (14.6), one compound of the 191 compounds in the Nagata data set was left out, and that one was thiophene, which residual value was far away from the rest.

The statistics of the final equation seem not to be all that bad taking into account the scatter of the data by Nagata, as mentioned, and the fact that the two data set have been obtained using totally different methods. The statistics of Eq. (14.6) are a little

better than those of Eq. (14.4), even though there are more data points in Eq. (14.6). The coefficients of the Abraham descriptors in Eq. (14.4) and in Eq. (14.6) differ somewhat, but the difference is within statistical error. For the coefficients e, s, a and b in Eq. (14.4) and Eq. (14.6) the error of the coefficient is about 0.4 unit, and for the l-coefficient the error is 0.05 unit.

14.2.3 Incorporation of the Hellman and Small data set

A set of odour detection thresholds for petrochemicals has been obtained by Hellman and Small (HS) [19] using a standard protocol. Although the ODT values were obtained many years ago, the data set includes several types of compound not present in the Nagata and Cometto-Muñiz and co-workers data sets, and so it seemed of interest to see if this set of ODT values could also be scaled to the Nagata set. We simply used the Hellman and Small data set (Table 14.8) as such, and incorporated a new descriptor in order to adjust this data set to the Nagata set. The descriptor **HS** = 1 for the HS set of log (1/ODT) values and **HS** = 0 for all other values. We did not include two aldehydes and two carboxylic acids in the Hellman and Small data set because there are not enough data points here to obtain indicator variables, and in any case we did not wish the resulting equation to be further complicated. In this case, one of the compounds of the Hellman and Small data set, i.e. carbitol acetate, had to be left out of the final set given a total data set of 323, because of its big residual. The equation is,

$$\begin{aligned}\text{Log (1/ODT)} = &- 1.491 + 0.298\,\mathbf{E} + 0.668\,\mathbf{S} + 1.107\,\mathbf{A} + 1.183\,\mathbf{B} + 0.576\,\mathbf{L} \\ &+ 4.068\,\mathbf{M} + 1.882\,\mathbf{AL} + 1.442\,\mathbf{AC} + 1.357\,\mathbf{UE} - 2.329\,\mathbf{C} \\ &+ 3.003\,\mathbf{CAC} + 1.652\,\mathbf{CAL} - 0.977\,\mathbf{HS}\end{aligned} \qquad (14.7)$$

$N = 323$, $R^2 = 0.743$, SD = 0.852, F = 68.74

The statistics of Eq. (14.7) are not quite as good as those of Eq. (14.6), but are comparable to those of Eq. (14.4). Of course, Eq. (14.7) is based on more data. Eq. (14.9) might be used in the prediction of further values of log (1/ODT) on the Nagata scale. It is not necessary to include the independent variables used to scale the Cometto-Muñiz and co-workers and the Hellman and Small ODT values. Hence the so final suggested equation to obtain odour detection thresholds on the Nagata scale is:

$$\text{Log (1/ODT)} = -1.491 + 0.298\,\mathbf{E} + 0.668\,\mathbf{S} + 1.107\,\mathbf{A} + 1.183\,\mathbf{B} + 0.576\,\mathbf{L} + 4.068\,\mathbf{M} + 1.882\,\mathbf{AL} + 1.442\,\mathbf{AC} + 1.357\,\mathbf{UE} \quad (14.8)$$

Table 14.8
Hellman and Small values of odour detection thresholds used to obtain Eq 14.9

Substance	log (1/ODT)	Substance	log (1/ODT)
ethene[a]	-2.415	diisopropylamine	0.886
propene	-1.352	dibutylamine[a]	1.097
buta-1,3-diene[a]	0.347	propylenediamine	0.136
dicyclopentadiene	-0.044	ethylene diamine[a]	0.229
5-ethylidene-2-norbornene[a]	1.699	methanol	-0.629
1,2-dichloroethane	-0.778	propan-2-ol[a]	-0.505
propylene dichloride[a]	0.110	butan-1-ol	0.523
fluorotrichloromethane	-0.699	2-methylpropan-1-ol[a]	0.167
trichlorotrifluoroethane[a]	0.091	butan-2-ol	0.921
diisopropylether	1.770	pentan-1-ol[a]	0.921
dibutylether[a]	1.155	pentanol	0.678
ethylene oxide	-2.415	2-methylbutan-1-ol[a]	1.398
1,2-propylene oxide[a]	-0.996	hexan-1-ol	2.000
1,2-butylene oxide	1.155	2-methylpentan-1-ol[a]	1.620
1,4-dioxane[a]	0.097	4-methylpentan-2-ol	0.481
propanone	-1.301	2-ethylbutan-1-ol[a]	1.155
butanone[a]	-0.301	2-ethylhexan-1-ol	1.125
4-methylpentan-2-one	1.000	diisobutyl carbinol[a]	-0.157
5-methylhexan-2-one[a]	1.921	isodecanol	-0.198
cyclohexanone	0.921	2-butoxyethanol[a]	1.000
2-methylpent-2-ene-4-one[a]	1.770	isobutyl cellosolve	-0.030
isophorone	0.699	diacetone alcohol[a]	-0.012
2,4-pentanedione[a]	0.073	methyl ethanolamine	0.161
ethyl acetate	-0.799	diethylethanolamine[a]	-0.046
propyl acetate[a]	1.301	toluene	0.770
isopropyl acetate	0.310	isopropylbenzene[a]	2.097
butyl acetate[a]	2.222	styrene	1.301
isobutyl acetate	0.456	α-methylstyrene[a]	1.284

Table 14.8 *(continued)*

Substance	log (1/ODT)	Substance	log (1/ODT)
2-ethylhexyl acetate[a]	1.000	styrene oxide	0.023
vinyl acetate	0.921	acetophenone[a]	0.523
2-methoxyethyl acetate[a]	0.469	1,3-dioxolane	0.268
2-ethoxyethylacetate	1.252	2-methylpyridine[a]	1.854
butyl cellosolve acetate[a]	-0.136	2-methyl-5-ethylpyridine	-0.041
ethylene diacetate[a]	1.032	morpholine[a]	0.141
isopropylamine	0.678	N-ethylmorpholine	-0.002
butylamine[a]	1.097	2-ethyl-1-hexyl acrylate	1.137
diethylamine	1.699	methyl methacrylate[a]	1.301
dipropylamine[a]	1.699		

[a]Part of the training set used for Eq. 14.11; see also Tables 14.2 and 14.8.

A plot of calculated values of log (1/ODT) on Eq. (14.7) against the observed values is shown in Figure 14.6. The three sets of experimental values are randomly distributed around the line of identity, showing that the indicator variables **C** and **HS** do indeed bring the Cometto-Muñiz and co-workers and the Hellman and Small odour detection thresholds onto the same scale as the Nagata thresholds. The value of using all three sets can be seen by the very wide range of the experimental log (1/ODT) values on the Nagata scale – almost 81/2 log units.

Although the Hellman and Small data set includes some unusual compounds, such as 5-ethylidene-2-norbornene, the chemical space of the three data sets is not very different. We can identify chemical space in terms of the five Abraham descriptors in Eq. (14.7). A principal component analysis of the values of the five descriptors yields five orthogonal PCs that contain all the information of the five descriptors. The first two PCs account for 62% of the total information, and a plot of the scores of PC2 against PC1 will indicate the chemical space in two dimensions, see Figure 14.7. There is quite an overlap of chemical space of the three data sets, but Figure 14.7 still shows how the chemical space of the Nagata data set can be expanded by incorporation of the other two data sets.

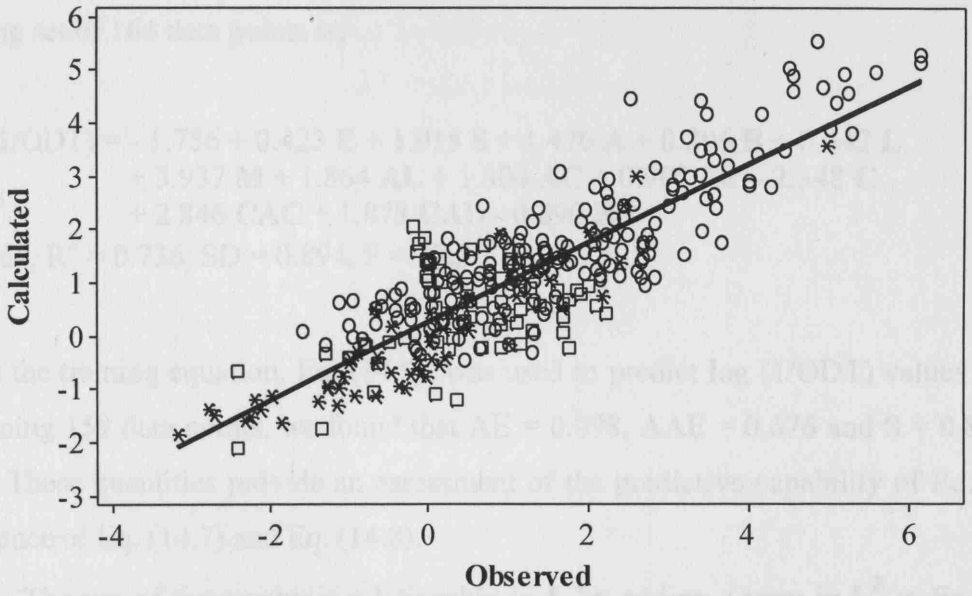


Figure 14.6. A plot of calculated values of log (1/ODT) on Eq. (14.7) against calculated values. ○ Nagata data; * Cometto-Muñiz and co-workers data; □ Hellman and Small data.

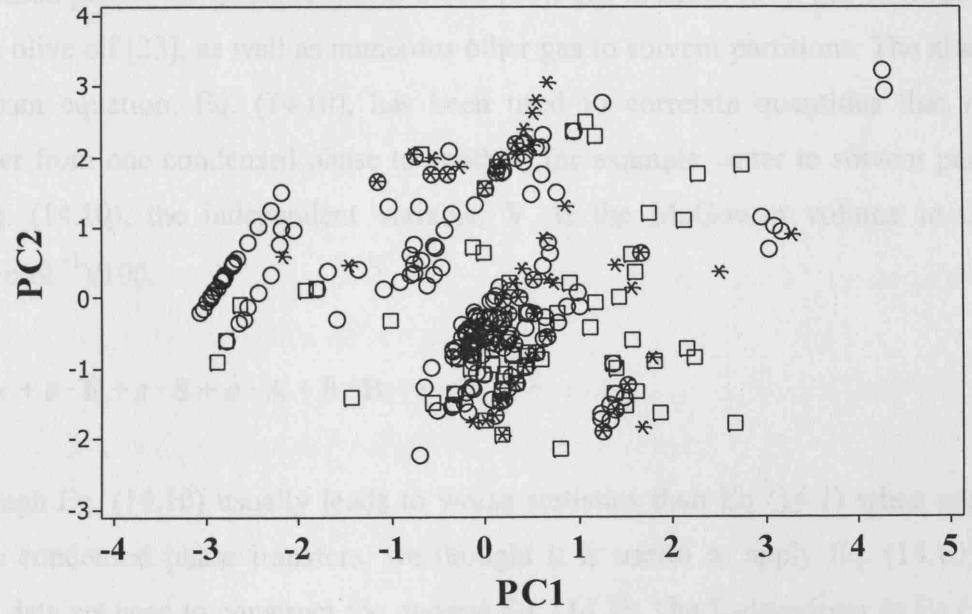


Figure 14.7. A plot of the scores of PC2 against the scores of PC1 for the five Abraham descriptors in Eq. (14.7) for all 323 data points. ○ Nagata data; * Cometto-Muñiz and co-workers data; □ Hellman and Small data.

An estimate of the predictive capability of Eq. (14.7) can be obtained by dividing the 323 ODT values into a training set and a test set. It is been arranged that all three data sets were included in the training set, and that the training set contained representative

samples of the various compounds with indicator variables. The equation for the training set of 164 data points is,

$$\begin{aligned}\text{Log (1/ODT)} = &- 1.756 + 0.423\,\mathbf{E} + 1.018\,\mathbf{S} + 1.476\,\mathbf{A} + 0.806\,\mathbf{B} + 0.642\,\mathbf{L} \\ &+ 3.937\,\mathbf{M} + 1.864\,\mathbf{AL} + 1.304\,\mathbf{AC} + 0.908\,\mathbf{ES} - 2.348\,\mathbf{C} \\ &+ 2.846\,\mathbf{CAC} + 1.878\,\mathbf{CAL} - 0.890\,\mathbf{HS}\end{aligned} \qquad (14.9)$$

$N = 164$, $R^2 = 0.736$, SD = 0.894, F = 32.2

When the training equation, Eq. (14.9), was used to predict log (1/ODT) values for the remaining 159 data points, we found that AE = 0.098, AAE = 0.676 and S = 0.849 log units. These quantities provide an assessment of the predictive capability of Eq. (14.9) and hence of Eq. (14.7) and Eq. (14.8).

The use of the parabolic relationship in **L** by adding a term in $\mathbf{L}^2$ to Eq. (14.7) was again investigated, but the resulting equation was no better than Eq. (14.7). An alternative equation, Eq. (14.10), was also investigated. The general Eq. (14.1) has invariably been used to correlate quantities that refer to transfer from the gas phase to a condensed phase, for example gas to blood [20], gas to brain [21], gas to muscle [22], gas to olive oil [23], as well as numerous other gas to solvent partitions. The alternative Abraham equation, Eq. (14.10), has been used to correlate quantities that refer to transfer from one condensed phase to another, for example water to solvent partitions. In Eq. (14.10), the independent variable, **V**, is the McGowan volume in units of $(cm^3 \cdot mol^{-1})/100$.

$$SP = c + e \cdot \mathbf{E} + s \cdot \mathbf{S} + a \cdot \mathbf{A} + b \cdot \mathbf{B} + v \cdot \mathbf{V} \qquad (14.10)$$

Although Eq. (14.10) usually leads to worse statistics than Eq (14.1) when applied to gas to condensed phase transfers, we thought it is useful to apply Eq. (14.10) to the entire data set used to construct the general Eq. (14.7). The **L**-descriptor in Eq (14.1) is usually obtained experimentally from data on gas chromatographic retention times [24] or can be estimated from fragment-based schemes [25,26]. However, there is no need even to estimate **V**, because it is specifically defined in terms of atom and bond contributions [27]. All that is required to calculate **V** is a knowledge of the molecular formula and a count of the number of bonds, B_n. The latter can be obtained trivially from the algorithm of Abraham [28]: $B_n = N_a - 1 - R$ where Na is the total number of

atoms in the molecule and R is the number of rings. There is thus an advantage of Eq. (14.10) over Eq. (14.1) in that one less descriptor needs to be determined or estimated. When we applied Eq. (14.10) to the 323 odour detection thresholds we obtained Eq. (14.11).

$$\begin{aligned}\text{Log (1/ODT)} = &- 2.021 + 1.083\,\mathbf{E} + 1.093\,\mathbf{S} + 1.421\,\mathbf{A} + 1.176\,\mathbf{B} + 2.062\,\mathbf{V} \\ &+ 4.148\,\mathbf{M} + 1.892\,\mathbf{AL} + 1.384\,\mathbf{AC} + 1.326\,\mathbf{UE} - 2.365\,\mathbf{C} \\ &+ 2.980\,\mathbf{CAC} + 1.743\,\mathbf{CAL} - 0.974\,\mathbf{HS}\end{aligned} \qquad (14.11)$$

$N = 323$, $R^2 = 0.742$, SD = 0.853, F = 68.5

Rather surprisingly, Eq. (14.11) is statistically equivalent to Eq. (14.7). The coefficients of the Abraham descriptors are not the same in Eq. (14.11) as in Eq. (14.7), because **V** and **L** encode somewhat different chemical information. Hence in comparison of coefficients for various gas to condensed phase processes Eq. (14.7) (or Eq. (14.8)) should be used. However, as a practical equation for the prediction of further values of odor detection thresholds on the Nagata scale, we recommend Eq. (14.12), the truncated version of Eq. (14.11).

$$\begin{aligned}\text{Log (1/ODT)} = &- 2.021 + 1.083\,\mathbf{E} + 1.093\,\mathbf{S} + 1.421\,\mathbf{A} + 1.176\,\mathbf{B} + 2.062\,\mathbf{V} \\ &+ 4.148\,\mathbf{M} + 1.892\,\mathbf{AL} + 1.384\,\mathbf{AC} + 1.326\,\mathbf{ES}\end{aligned} \qquad (14.12)$$

Eq. (14.12), as with Eqs. (14.7)-(14.9) and Eq. (14.11), should lead to predictions of log (1/ODT) to within 0.68 (AAE) or 0.85 (S) log units. Of course, the same *caveats* with respect to reactive compounds will apply to Eq. (14.12) as to Eq. (14.8).

The principal component analysis, the regression equations and the various calculations were all carried out using Minitab [29] software.

14.3. REFERENCES

1. J. E. Cometto-Muñiz, W. S. Cain, M. H. Abraham, Exp. Brain Res. 118 (1998) 180
2. J. E. Cometto-Muñiz, W. S. Cain, Indoor Air 4 (1994) 140
3. R. Schmidt, W. S. Cain, (2006) The credibility of measured odor thresholds, Presented at the 28th Annual Meeting of the Association for Chemoreception Sciences, Sarasota, FL, April 28, 2006

4. Y. Nagata (2003) Measurement of odour threshold by triangle odour bag method. In *Odour measurement review*. Office of Odour, Noise and Vibration. Environmental Management Bureau, Ministry of Environment, Tokyo, pp. 118-127
5. J. E. Cometto-Muñiz, Physicochemical basis for odor and irritation potency of VOCs in *Indoor Air Quality Handbook*, J.D. Spengler, J. Samet, and J. F. McCarthy (eds.), Mcgraw-Hill, New York, 2001, p. 20.1
6. M. H. Abraham, J. M. R. Gola, J. E. Cometto-Muñiz, W.S. Cain, Chem. Senses 27 (2002) 95
7. American Industrial Hygiene Association (AIHA) (1989) Odour thresholds for chemicals with established occupational standards, Akron: AIHA.
8. M. Devos, F. Patte, J. Rouault, P. Laffort, L. J. Van Gemert (1990) Standardized human olfactory thresholds, IRL Press at Oxford University Press, Oxford.
9. Environmental Protection Agency (US EPA) (1992) *Reference guide to odour thresholds for hazardous air pollutants listed in the clean air act amendments of 1990*. EPA600/R-92/047, March 1992
10. J. D. Pierce, R. L. Doty, J. E. Amoore, Percept. Mot. Skills 82 (1996) 451-458
11. M. F. Dube, C. R. Green, Rec. Adv. Tob. Sci. 8 (1982) 42-102
12. M. H. Abraham, J. Andonian-Haftvan, G. S. Whiting, A. Leob, R. S. Taft, J. Chem. Soc., Perkin Trans. 2 (1994) 1777
13. M. H. Abraham, G. S. Whiting, P. W. Carr, H. Ouyang, J. Chem. Soc., Perkin Trans. 2 (1998) 1385
14. M. H. Abraham, J. A. Platts, A. Hersey, A. J. Leo, R. W. Taft, J. Pharm. Sci. 88 (1999) 670
15. M. H. Abraham, J. Le, W. E. Acree, Jr., Collect. Czech. Chem. Commun. 64 (1999) 1748
16. M. H. Abraham, P. K. Weathersby, J. Pharm. Sci. 83 (1994) 1450
17. Y. Alarie, M. Schaper, G. D. Nielsen, M. H. Abraham, Arch. Toxicol. 72 (1998) 125
18. Y. Alarie, G. D. Nielsen, M. H. Abraham, Pharmacol. Toxicol. 83 (1998) 270
19. T. M. Hellman, F. H. Small, J. Air Pollut. Control Assoc. 24 (1974) 979
20. M. H. Abraham, A. Ibrahim, W. E. Acree, Jr., Chem. Res. Toxicol. 18 (2005) 904
21. M. H. Abraham, A. Ibrahim, W. E. Acree, Jr., Eur. J. Med. Chem. 41 (2006) 494
22. M. H. Abraham, A. Ibrahim, W. E. Acree, Jr., Chem. Res. Toxicol. 19 (2006) 801
23. M. H. Abraham, A. Ibrahim, Chem. Inf. Model. 46 (2006) 1735

24. M. H. Abraham, A. Ibrahim, A. M. Zissimos, J. Chromatogr. A. 1037 (2004) 29
25. J. A. Platts, D. Butina, M. H. Abraham, A. J. Hersey, Chem. Inf. Comp. Sci. 39 (1999) 835
26. PharmaAlgorithms (2006) ADME Boxes, Version 3.0, PharmaAlgorithms Inc., 591 Indian Road, Toronto, ON M6P 2C4, Canada.
27. M. H. Abraham, J. C. McGowan, Chromatographia 23 (1987) 243-246
28. M.H. Abraham, Chem. Soc. Revs. 22 (1993) 73-83
29. Minitab, Version 14 (2003) Minitab Inc., Quality Plaza, 1829 Pine Hall road, State College, Pa 16801-3008, USA

Chapter 15 Concentration-Detection Functions for the Olfactory Detectability of Airborne Chemicals

15.0. INTRODUCTION

Advancements in molecular biology have helped to clarify the cellular and molecular events that take place in olfactory sensory neurons (OSN) when they are stimulated by an odorant [1]. The interaction of the stimulus with an olfactory receptor (OR) leads to a cascade of events that have been described in some detail and involve the transduction of the chemical message into an electrical message that is conducted first to the olfactory bulb, and then to higher levels of the olfactory pathway [2]. In humans, it has been estimated that there are approximately 388 genes coding for functional olfactory receptors and about 414 pseudogenes that do not code for functional receptors [3]. The present main line of thought about the workings of the olfactory system states that each OSN expresses one OR and that ORs form a continuous gradient across the olfactory epithelium (OE) [4]. The axons of OSN expressing the same receptor converge on the same glomeruli on the olfactory bulb [5].

At the cellular and molecular level of olfaction, considerable information has been gathered about the secondary message system cascade, see review in [6], whereas less is known about the primary interaction between odorants and receptors from a structure-activity point of view [7-11]. Among other factors, the reason for this could rest in that the secondary message pathways are analogous to those discovered earlier in vision [12]. Also, the large number of odorant receptors (and odorants), the challenge of expressing them in a suitable system [13,14], and their relative low ligand-binding affinities [10] have posed some difficulties for structure-activity studies at the molecular level. At the integrated or psychophysical olfactory level, there have been a number of attempts to establish odour structure-activity relationships (SARs), see reviews in [15,16], but most of them, first, focused on the suprathreshold level, and, second, lacked mechanistic significance. In contrast, few studies addressed odour SARs at the threshold level [17] and, furthermore, these also rested on a mechanistic basis [18,19].

In the present study concentration-detection, i.e., detectability, functions for the human olfactory detection of 18 chemical vapours of diverse chemical structures measured by Cometto-Muñiz and co-workers in the past few years have been obtained

[20]. The functions cover the complete peri-threshold range from chance to almost perfect detection and provide much more information than what a single "threshold value" can convey. In this chapter, these odour functions have been modeled with a uniform approach: a simplified sigmoid equation (see below), and have been correlated with a set of up to five physicochemical descriptors taken from the general solvation equation of mechanistic significance that has quite successfully described the simpler "odour threshold" psychophysical outcome, as described in the last two chapters of this work.

15.1. MATERIALS AND METHODS

15.1.1 Subjects and stimuli.

All participants were normosmics (i.e., had a normal sense of smell) as determined by a clinical olfactory test [21]. Table 15.1 describes the characteristics of all subjects tested. In this table the chemicals tested, the delivery techniques, the number of subjects (#S's), their average age (±SD) and range, the gender distribution (Females(F)/Males(M)), the number of trials per subject/concentration, and the total number of trials per concentration, can also be found. A total of 18 stimuli were tested. Delivery techniques include: squeeze bottle (SB), glass vessel (GV), and vapour delivery device (VDD), see below [20].

Table 15.1
Characteristics of all subjects tested [20]

Chemical Stimulus	Delivery Technique	# S's	Average Age (±SD)	Age Range	F/M	Trials per subject	Total trials
1. ethanol	VDD	4	29(±14)	21-49	2/2	≥15	61
1. ethanol	GV	19		18-43		30-40	570
2. 1-butanol	SB	4	36(±13)	24-54	3/1	16	64
2. 1-butanol	VDD	7	32(±15)	22-61	2/5	10	70
3. 2-heptanone	SB	4	36(±13)	24-54	3/1	16	64
4. hexanoic acid	VDD	4	40(±16)	24-61	0/4	28	112
5. glutaraldehyde	VDD	40	22(±3)	18-35	40/0	28	1120
6. butyl acetate	SB	4	37(±14)	25-56	3/1	16	64

Table 15.1 *(continued)*

Chemical Stimulus	Delivery Technique	# S's	Average Age (±SD)	Age Range	F/M	Trials per subject	Total trials
7. ethyl propanoate	GV	22	26(±10)	18-50	10/12	≥27	610
8. ethyl butanoate	VDD	4	22(±2)	20-25	2/2	100	400
9. ethyl heptanoate	GV	22	26(±10)	18-50	10/12	≥29	658
10. benzene	VDD	3	37(±22)	24-63	1/2	16	48
11. toluene	SB	4	37(±14)	25-56	3/1	16	64
12. naphthalene	GV	20	25(±6)	19-40	10/10	28	560
13. 1-methyl naphthalene	GV	20	25(±6)	19-40	10/10	28	560
14. 2-methyl naphthalene	GV	20	25(±6)	19-40	9/11	28	560
15. D-limonene	VDD	11	25(±5)	20-33	7/4	30	330
16. chloropicrin	VDD	43	23(±4)	19-34	18/25	30	1290
17. ozone	VDD	11	26(±5)	20-33	6/7	30	330
18. TXIB	GV	19		18-43		30-40	570

Presentation of stimuli [22] involved a dynamic system via a vapour delivery device (VDD) [23], and/or a static system via squeeze bottles (SB) and/or glass vessels (GV) [23] (see Table 15.1). For odour stimulation with a static system, SB and GV ended, respectively, in a single spout or two nosepieces from which the subjects sniffed the headspace of the container [24]. (When using SB, subjects were instructed to squeeze with approximately equal strength on all trials.) For odour stimulation with the dynamic VDD system, the linear velocity of stimulus and blanks (i.e., carbon-filtered air) was ≈13 cm/sec, similar to that found in a typical indoor environment [25,26], even when the volume flow (40 L/min) was high enough to fully accommodate the most forcible instantaneous human odour sniffs [27,28]. This was achieved by delivering the sample from specially designed glass cones where the participant exposed the nose [23].

15.1.2 Procedure.

All odour testing involved using a two- or three-alternative forced-choice procedure between stimulus and blanks [29]. For static delivery, blanks comprised the headspace

above mineral oil (light, FCC) carried by either nitrogen or air. For dynamic delivery, blanks comprised carbon-filtered air.

15.1.3 Data analysis.

The outcome is summarized in terms of odour detection probability, i.e., detectability, as a function of stimulus vapour concentration (log ppm). Detectability was corrected for chance according to [29]:

$$P = [m \cdot p(c) - 1] / (m - 1) \tag{15.1}$$

where P = detection probability corrected for chance, m = number of choices in the forced-choice procedure (i.e., 2 or 3), and p(c) = proportion correct (i.e., number of correct trials / total number of trials).

Odour concentration-detection (also called psychometric or detectability) functions were modeled by a simplified sigmoid equation of the form:

$$y = 1 / (1 + e^{[-(x-C)/D]}) \tag{15.2}$$

where y = detectability (P) as defined in Eq. (15.1), x = vapor concentration of the chemical stimulus (in log ppm by volume), and C and D are constants. Note that C represents the concentration of the stimulus (i.e., x) when y = 0.5, that is, when detectability is half way (i.e., P =0.5) between chance detection (i.e., P = 0.0) and perfect detection (i.e., P = 1.0). In turn, the value of constant D governs the steepness of the detectability function.

15.2. RESULTS

Figure 15.1 presents odour functions for the 18 chemicals. Table 15.2 lists the values for the constants C and D, with their respective standard errors, and the values for Chi-square and the coefficient of determination, R^2, of the functions. Both the figure and the table show that Eq. (15.2) provides an excellent description of the experimental data. As mentioned under Data analysis, the value of C represents the concentration of the chemical producing a level of detectability situated half way between chance

detection and perfect detection. Such concentration is commonly, although not always, taken as the threshold concentration. Use of alternative definitions of thresholds and psychophysical procedures would produce different absolute values for the threshold of any particular chemical (cf. [30]). Nevertheless, under standardize conditions of measurement, the relative chemosensory potency of chemicals (as gauged, for example, by a threshold value or by a psychometric function) should remain largely unaltered. A convenient approach to assess relative potency of vapours is to test homologous chemical series. It is well established that olfactory potency increases (i.e., thresholds decline) along such series [31]. It was of interest to confirm if the psychometric functions presented here could also capture this trend.

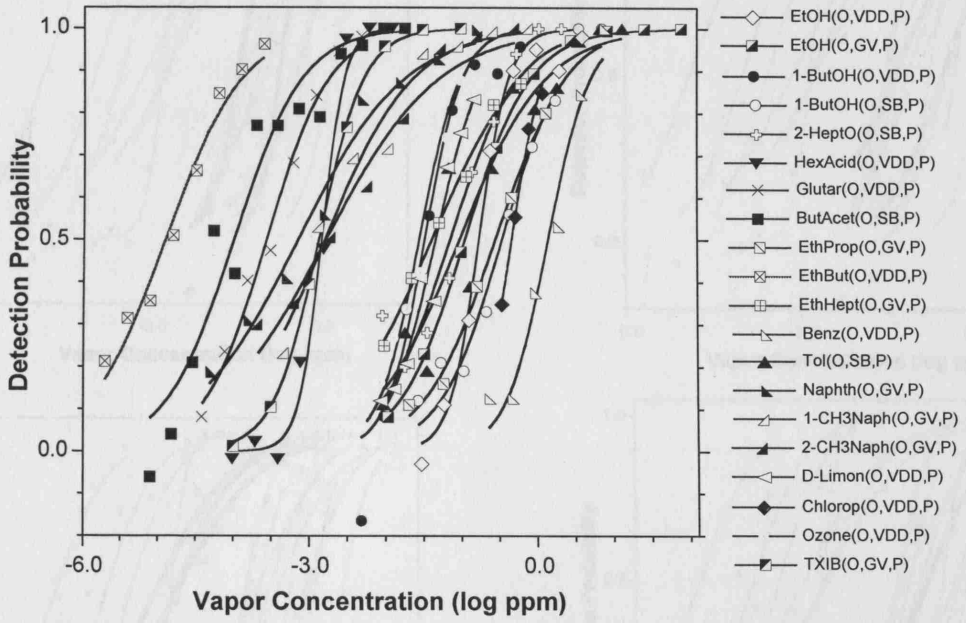


Figure 15.1. Odour psychometric functions modeled by Eq. (15.2) for chemicals and techniques listed in Table 15.2.

Odour potency across homologs tested under identical method and procedure (Table 15.2) for the following series can be compared: 1) n-alcohols tested via VDD (ethanol and 1-butanol) and 2) esters tested via GV (ethyl propanoate and ethyl heptanoate). In both cases, the functions clearly show the expected trend: the larger homolog has a higher potency, as reflected by curves shifted to the left, i.e., towards lower concentrations (Figure 15.2). Note that six-carbon esters, i.e., ethyl butanoate and butyl acetate, are more potent than expected from the other esters. This outcome cannot be solely explained, as in the case of ethyl butanoate, by use of the optimized VDD system

(which has produced low odour thresholds across stimuli [23], since the result still holds, as in the case of butyl acetate, when using the simpler SB system (which typically has produced high odour thresholds across stimuli [32]). Regarding alkylbenzenes one could also consider a tentative comparison of odour potency, despite the fact that benzene was tested via VDD whereas toluene was tested via SB. In spite of the presentation differences favouring benzene, toluene still shows higher potency than benzene, confirming the expected trend (Figure 15.2). In the case of the bulkier molecule naphthalene, addition of a methyl group in position 1 or 2 does not displace the odour function to any great extent (Figure 15.2).

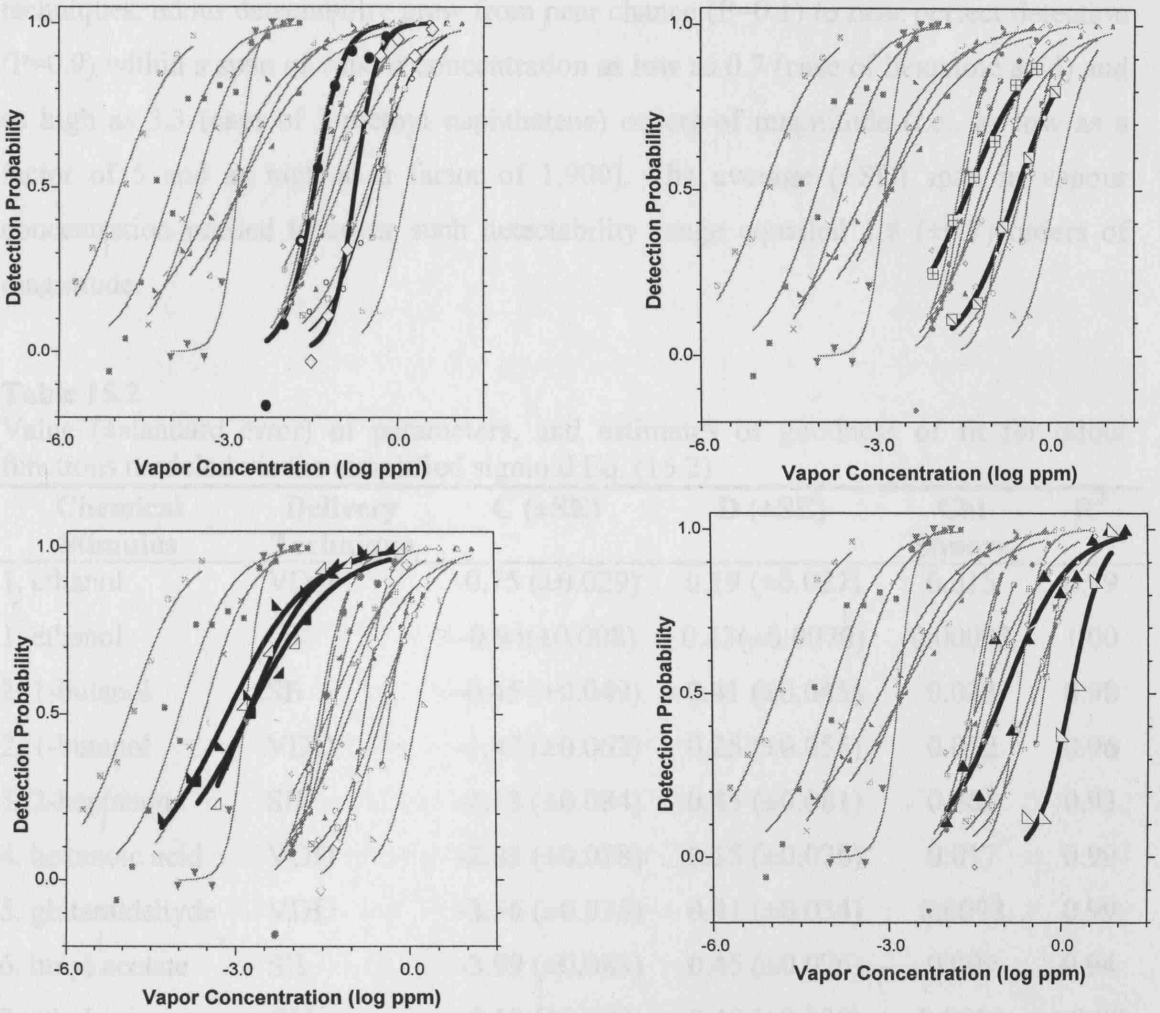


Figure 15.2. Illustrating, against the background of all odour functions, how homologs gain olfactory potency, i.e., functions shift towards lower concentrations, with increasing carbon chain length. Upper left: Ethanol (empty diamonds) and 1-butanol (filled circles). Upper right: Ethyl propanoate (squares with a diagonal) and ethyl heptanoate (squares with a cross). Lower left: Benzene (empty triangles) and toluene (filled triangles). Lower right: In contrast, the bulkier stimuli naphthalene (filled rectangle triangles facing right), 1-methyl naphthalene (empty rectangle triangles facing left), and 2-methyl naphthalene (filled rectangle triangles facing left) produced largely overlapping functions.

Regarding the steepness of the odour functions, governed by the constant D on Eq. (15.2), there are not radical differences among distinct chemicals or clear trends among related homologs within this varied set of stimuli (Table 15.2). Note, however, that, except for benzene, the aromatic compounds (toluene and all three naphthalenes) tend to produce shallower functions, i.e., functions with a higher value of D, than the rest of the stimuli. Excluding benzene, the remaining four aromatics show an average (±SD) D = 0.65 (±0.09), whereas all other compounds show an average D = 0.36 (±0.13). Averaging across all odour functions produces D = 0.40 (±0.18).

An analysis of the functions in Figure 15.1 reveals that, across chemicals and techniques, odour detectability grew from near chance (P=0.1) to near perfect detection (P=0.9) within a span of vapour concentration as low as 0.7 (case of hexanoic acid) and as high as 3.3 (case of 2-methyl naphthalene) orders of magnitude (i.e., as low as a factor of 5 and as high as a factor of 1,900). The average (±SD) span in vapour concentration needed to cover such detectability range equalled 1.8 (±0.7) orders of magnitude.

Table 15.2
Value (±standard error) of parameters, and estimates of goodness of fit for odour functions modeled via the simplified sigmoid Eq. (15.2)

Chemical Stimulus	Delivery Technique	C (±SE)	D (±SE)	Chi Square	R^2
1. ethanol	VDD	-0.75 (±0.029)	0.19 (±0.027)	0.015	0.99
1. ethanol	GV	-0.94(±0.008)	0.43(±0.0070)	0.00033	1.00
2. 1-butanol	SB	-0.45 (±0.049)	0.41 (±0.045)	0.023	0.98
2. 1-butanol	VDD	-1.47 (±0.062)	0.25 (±0.055)	0.052	0.96
3. 2-heptanone	SB	-1.13 (±0.084)	0.45 (±0.081)	0.061	0.93
4. hexanoic acid	VDD	-2.83 (±0.028)	0.15 (±0.026)	0.017	0.99
5. glutaraldehyde	VDD	-3.56 (±0.035)	0.41 (±0.034)	0.0098	0.99
6. butyl acetate	SB	-3.99 (±0.083)	0.45 (±0.076)	0.099	0.94
7. ethyl propanoate	GV	-0.53 (±0.029)	0.48 (±0.030)	0.0019	>0.99
8. ethyl butanoate	VDD	-4.91 (±0.044)	0.50 (±0.045)	0.012	0.98
9. ethyl heptanoate	GV	-1.37 (±0.028)	0.60 (±0.033)	0.0020	>0.99
10. benzene	VDD	0.16(±0.047)	0.28 (±0.043)	0.017	0.97

Table 15.2 *(continued)*

Chemical Stimulus	Delivery Technique	C (±SE)	D (±SE)	Chi Square	R^2
11. toluene	SB	-1.01 (±0.110)	0.57 (±0.107)	0.119	0.90
12. naphthalene	GV	-3.14 (±0.064)	0.70 (±0.063)	0.011	0.98
13. 1-methyl naphthalene	GV	-2.78 (±0.077)	0.58 (±0.081)	0.018	0.96
14. 2-methyl naphthalene	GV	-2.77 (±0.080)	0.75 (±0.083)	0.016	0.97
15. D-limonene	VDD	-1.32 (±0.038)	0.31 (±0.039)	0.024	0.96
16. chloropicrin	VDD	-0.32 (±0.012)	0.21 (±0.014)	0.00090	>0.99
17. ozone	VDD	-1.54 (±0.040)	0.27 (±0.039)	0.033	0.95
18. TXIB	GV	-2.87 (±0.007)	0.30 (±0.007)	0.00032	1.00

15.3. DISCUSSION

Previously, the very general equation for the correlation of processes entailing the transfer of chemical vapours form the gas phase to a condensed (bio)phase was applied to olfaction [18]. This equation takes the following form [33,34]:

$$SP = c + e \cdot \mathbf{E} + s \cdot \mathbf{S} + a \cdot \mathbf{A} + b \cdot \mathbf{B} + l \cdot \mathbf{L} \qquad (15.3)$$

As mentioned in previous chapters, in a chemosensory context, SP stands for the respective reciprocals of: an odour detection threshold as log(1/ODT), a nasal chemesthetic (i.e., nasal pungency or irritation) threshold as log(1/NPT), or an ocular chemesthetic (i.e., eye irritation) threshold as log(1/EIT), where all thresholds are expressed as log ppm by volume [18,35-37]. The bold letters in Eq. (15.3) represent physicochemical descriptors of volatile compounds as follows: **E** is an excess molar refraction, **S** is the dipolarity/polarizabiltiy, **A** and **B** are the overall or effective hydrogen-bond acidity and basicity, respectively, and **L** (log L^{16}) is defined through L^{16}, the volatile compound (i.e., the solute) Ostwald solubility coefficient on hexadecane at 298 K. The **L**-descriptor is itself a combination of two solute properties: (i) a general measure of solute size, and (ii) the ability of a solute to interact with a solvent phase (i.e., the receptor environment or biophase) through dispersion forces. The units of **E** are $cm^3/10$ whereas the other descriptors have no units because they are all derived from

the logarithm of an equilibrium constant. In turn, the coefficients *c*, *e*, *s*, *a*, *b*, and *l* are found by multiple linear regression analysis, Nevertheless, they are not simply fitting coefficients, because they reflect the complementary properties that must characterize the receptor biophase for an optimum interaction with the solute, that is, the dissolved odorant. In this way, the solvation Eq. (15.3) conveys mechanistic information on the interaction of the odorant (or solute) with the odorant-receptor environment (or solvent) as follows: The *e*-coefficient gives the tendency of the receptor biophase to interact with the odorant through polarizability-type interactions, mostly via electron pairs. The *s*-coefficient is a measure of the receptor biophase dipolarity/polarizability. The *a*-coefficient represents the complementary property to the odorant hydrogen-bond acidity and thus is a measure of the receptor biophase hydrogen-bond basicity. Likewise, the *b*-coefficient is a measure of the receptor biophase hydrogen-bond acidity. Finally, the *l*-coefficient is a measure of the hydrophobicity of the receptor biophase.

Eq. (15.3) has also been applied to numerous gas-solvent partitions [38-41], to other gas-biophase partitions including trigeminal chemesthesis [34,35], and to a very large number of gas chromatographic systems [42], so it is a well tried and tested equation. Eq. (15.3) is best used to model transport or transfer processes, that is, those in which either the distribution of a solute between biophases or the rate of transfer of a solute from one biophase to another forms the key step. These processes are selective, in that different odorants will have different equilibrium constants depending on their structure, but not specific, in those small changes in structure or small positional changes of functional groups have little effects on them. For odour thresholds, as shown previously, the selective transport of odorants account for 81% of the total effect, a percentage that increases to 86% with the addition to Eq. (15.3) of a parabolic term in L (an odorant size parameter) [18]. The remaining portion of the effect might be due to specific-receptor ligand effects.

The results of applying Eq. (15.3) not simply to odour thresholds but to the complete odour detectability functions presented here, as defined by the constants C and D, will now be presented.

As a first step, Eq. (15.3) was applied to the C and D constants listed in Table 15.2. The physicochemical descriptors of the corresponding compounds can be found in Table 15.3. In each case, several of the descriptors were not statistically significant, and the resulting equations are given as Eq. (15.4) and Eq. (15.5). Note that the C and D values for ethanol (GV) and TXIB (GV) were not used for this purpose. In Eq. (15.5) an

indicator variable, **VDD**, was introduced. This indicator variable takes the value **VDD** = 1 for compounds measured by the VDD method and **VDD** = 0 for compounds measured by any of the other methods. The indicator variable was not significant in Eq. (15.4).

$$C = 2.33 - 3.05\ \mathbf{B} - 0.75\ \mathbf{L} \qquad (15.4)$$

$N = 15$, $SD = 0.75$, $R^2 = 0.65$, $F = 10.99$

$$D = 0.49 + 0.13\ \mathbf{E} - 0.21\ \mathbf{A} - 0.23\ \mathbf{VDD} \qquad (15.5)$$

$N = 17$, $SD = 0.07$, $R^2 = 0.88$, $F = 31.04$

At first sight, the statistics of Eq. (15.4) and Eq. (15.5) are not particularly good, but the low number of compounds studied is probably the cause of this. However, the good agreement between psychometric functions obtained from C and D calculated from Eq. (15.4) and Eq. (15.5) and experimentally obtained psychometric functions shows that the calculated C and D values must be very reasonable. Indeed, this good agreement, as shown in Figure 15.3, is a more cogent indicator of the usefulness of Eq. (15.4) and Eq. (15.5) than any statistical parameter.

Table 15.3
Descriptors for the set of compounds used to build Eqs. (15.4) and (15.5)

Chemical Stimulus	E	S	A	B	L
ethanol	0.246	0.42	0.37	0.48	1.485
butan-1-ol	0.224	0.42	0.37	0.48	2.601
2-heptanone	0.123	0.68	0.00	0.51	3.760
hexanoic acid	0.174	0.63	0.62	0.44	3.697
glutaraldehyde	0.338	1.10	0.00	0.75	3.800
butyl acetate	0.071	0.60	0.00	0.45	3.353
ethyl propanoate	0.087	0.58	0.00	0.45	2.807
ethyl butanoate	0.068	0.58	0.00	0.45	3.271
ethyl heptanoate	0.027	0.58	0.00	0.45	4.733
benzene	0.610	0.52	0.00	0.14	2.786
toluene	0.601	0.52	0.00	0.14	3.325
naphthalene	1.340	0.92	0.00	0.20	5.161
1-methyl naphthalene	1.337	0.94	0.00	0.22	5.802

Table 15.3 *(continued)*

Chemical Stimulus	E	S	A	B	L
2-methyl naphthalene	1.304	0.81	0.00	0.25	5.617
D-limonene	0.488	0.28	0.00	0.21	4.725
chloropicrin	0.461	0.82	0.00	0.10	3.208
Ozone	0.000	0.10	0.09	0.00	0.039

In Figure 15.3 are given the experimentally observed psychometric functions side by side with those obtained using the calculated C and D values from Eq. (15.4) and Eq. (15.5). The continuous lines represent the functions according to the C and D constants in Table 15.2, and the dashed lines are the functions obtained from the calculated C and D values from Eqs. (15.4) and (15.5). In most cases the shape of the psychometric functions from the experimental and calculated C and D values is very similar, and the functions are very close to each other. In some cases the experimental and calculated functions even overlap, e.g. chloropicrin(O, VDD, P), 1-methylnaphthalene(O, GV, P), and 2-methylnaphthalene(O, GV, P). This suggest that it is possible to predict the entire psychometric function for an odorant simply by predicting the relevant C and D values from Eq. (15.4) and Eq. (15.5), using already known values of the descriptors.

On the other hand, in Figure 15.4 are shown the psychometric functions for three odorants, where the experimental and calculated functions are very different. The three outliers found are butyl acetate (SB), ethyl butanoate (VDD), and ozone (VDD). For the first two, their function steepness (constant D) can be described rather well but they are much more potent (constant C) than predicted. The same holds for the third outlier, ozone, but since ozone is a reactive chemical it is not likely to be able to fit the psychometric function by calculated C and D values.

It is not so obvious why butyl acetate (SB) and ethyl butanoate (VDD) are outliers. The different methods used to obtain the psychometric functions do not seem relevant, because the C and D values for butan-1-ol were obtained using two different methods, the SB method and the VDD method, and good fits were obtained in both cases. A plot of the experimental C values against the experimental D values is shown in Figure 15.5. In this plot, the three outliers are shown each surrounded by a circle. It seems likely that all three psychometric functions should have a D value smaller than 1 and a C value bigger than –4.

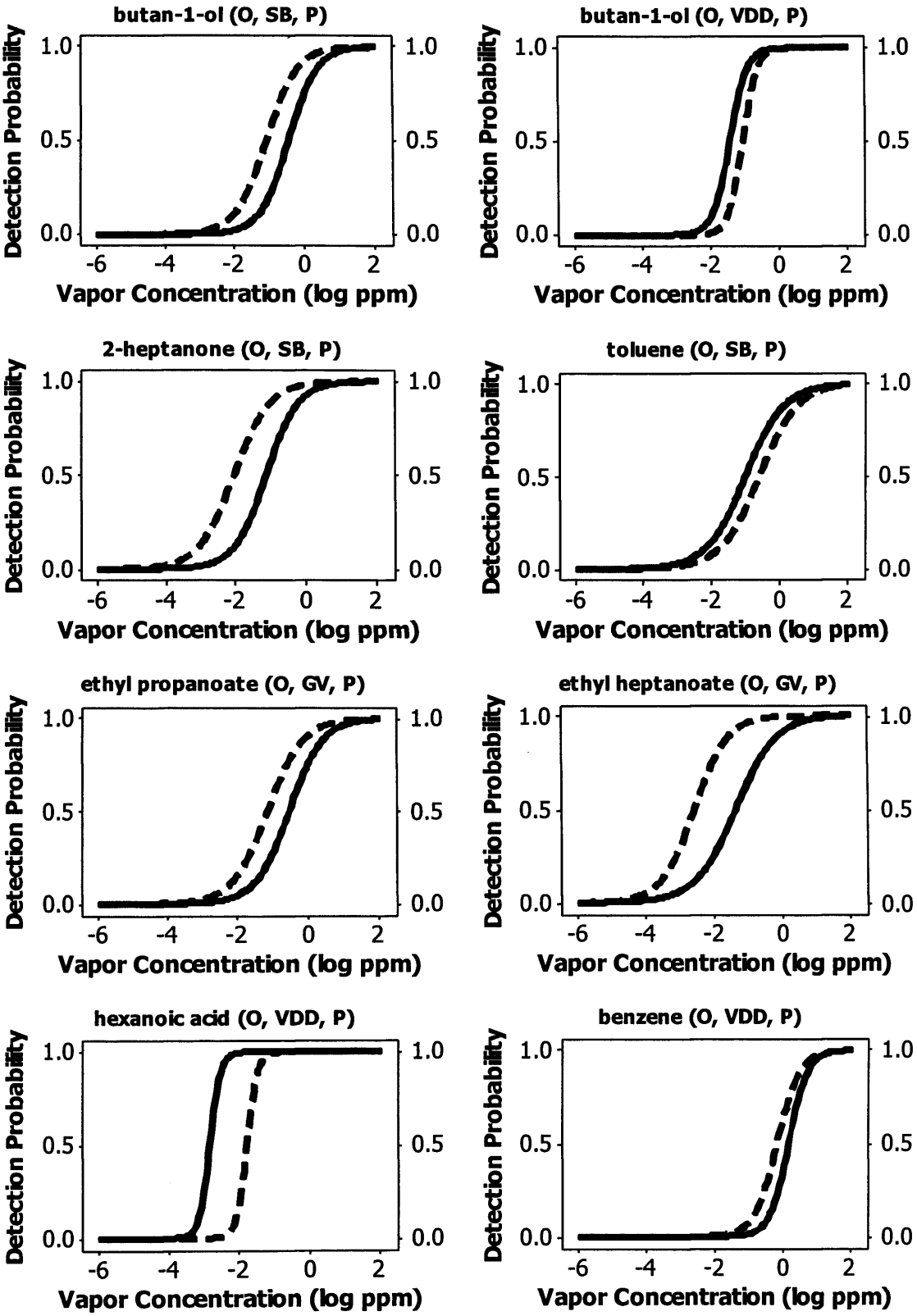


Figure 15.3. Comparison of the experimental (————) and calculated (– – – – –) odour psychometric functions for those chemicals used to obtain Eqs. (15.4) and (15.5).

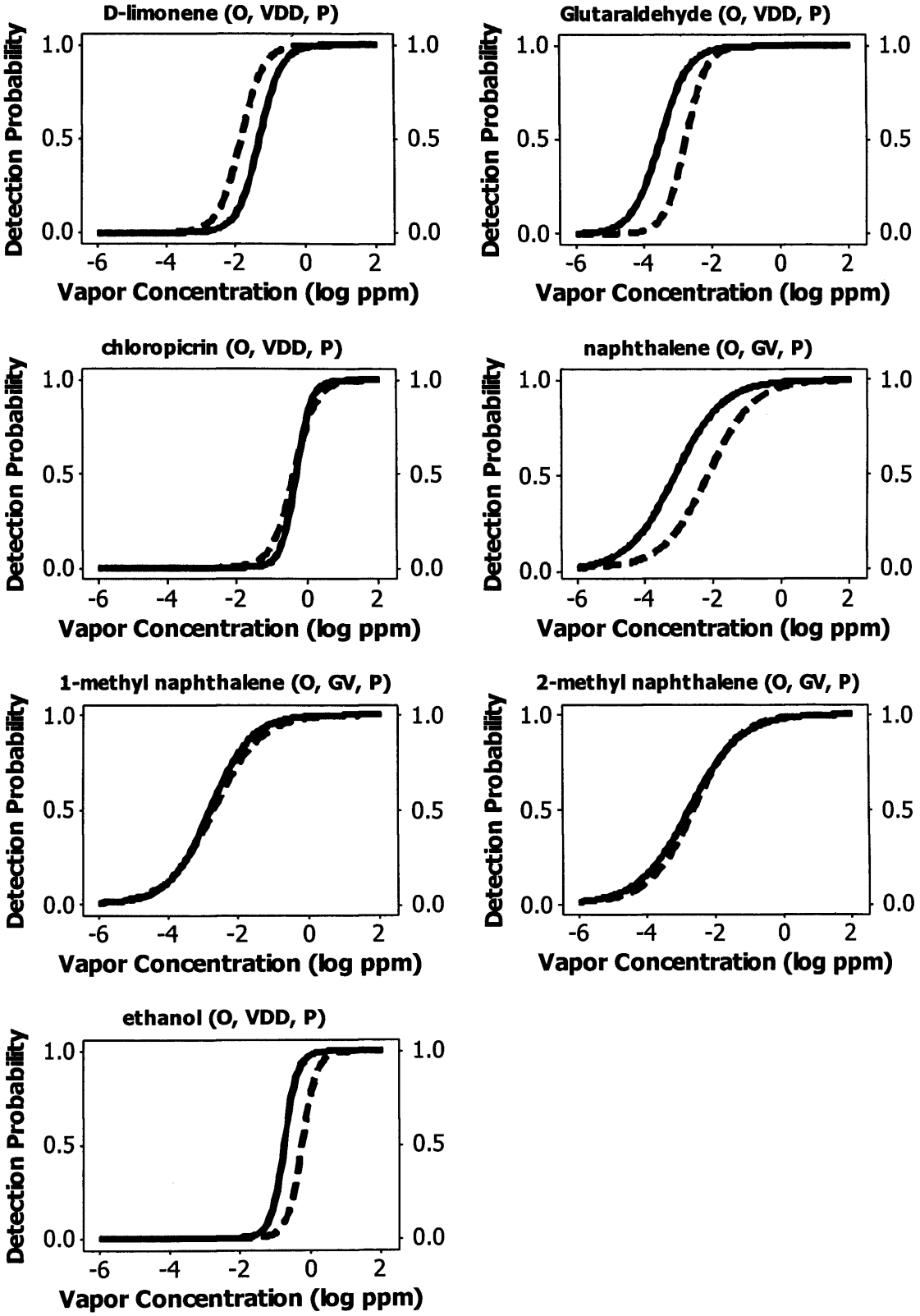


Figure 15.3. *(continued)*

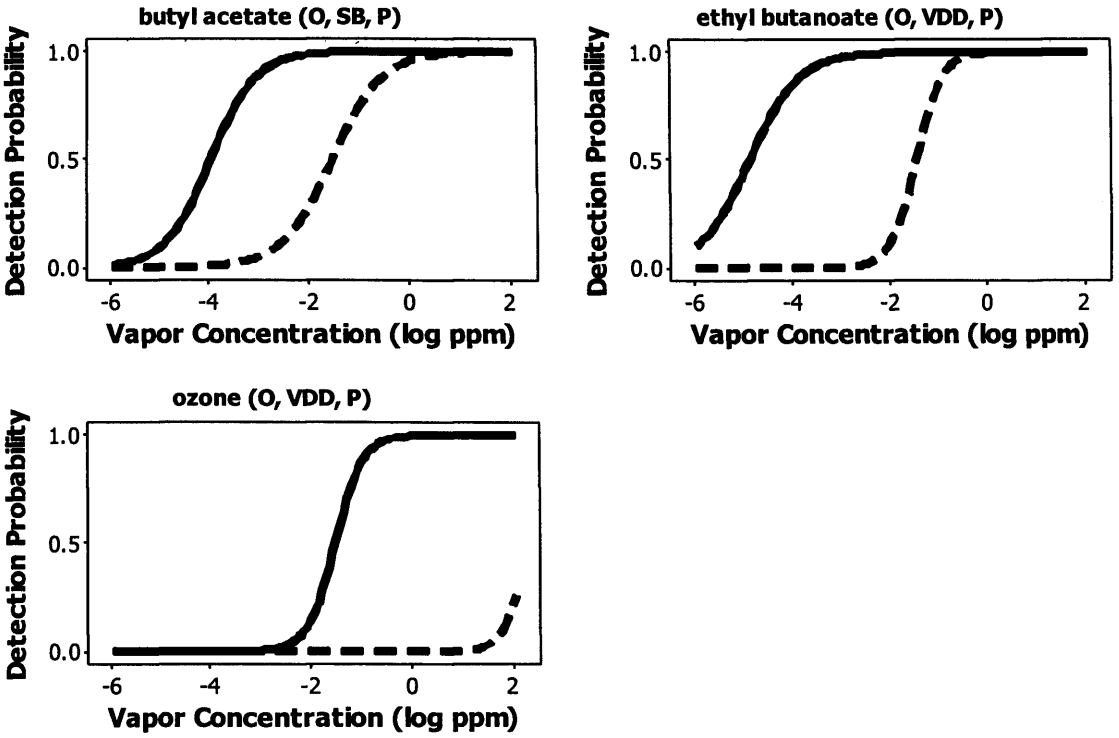


Figure 15.4. Odour psychometric functions for those chemicals found as outliers, experimental (————) and calculated (– – – – –).

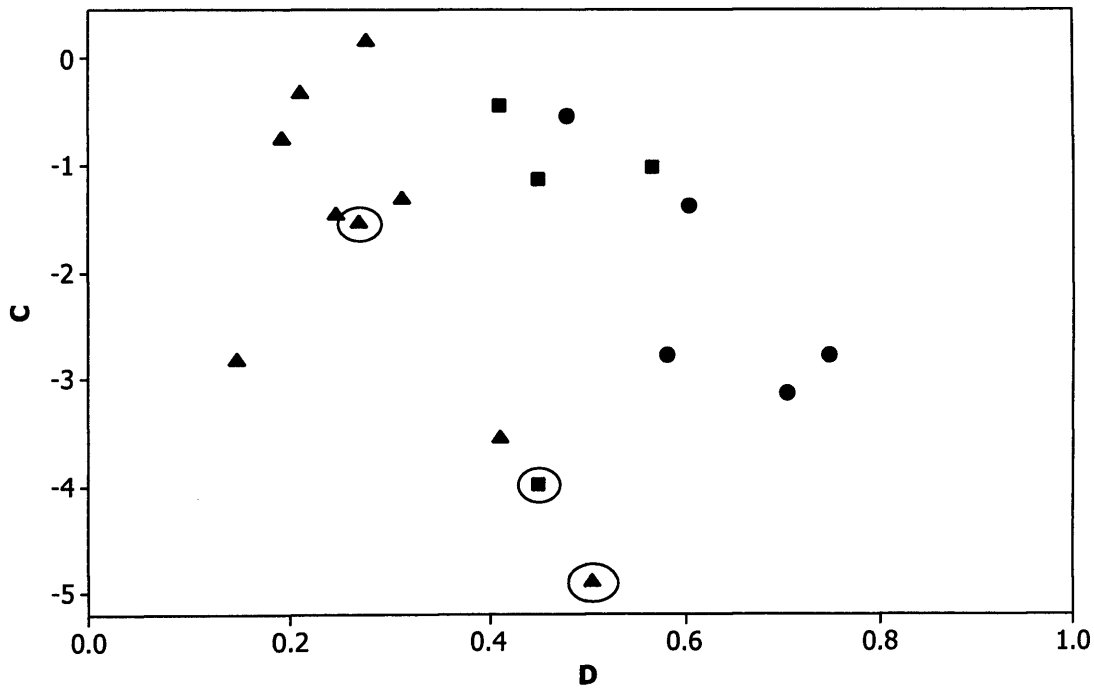


Figure 15.5. A plot of the experimental values of C against D. Surrounded by a circle are the outliers found during the regression analysis. Dynamic system via a vapour delivery device VDD (filled triangle), static system via squeeze bottles SB (filled squares), glass vessels GV (filled circles).

This is also supported by the fact that none of the calculated values for D is bigger than 1 and none of the calculated values for C is smaller than –4, see Table 15.4. Inspection of the experimental C values in Table 15.2 shows that those for butyl acetate (-3.99) and ethyl butanoate (-4.91) are not at all compatible with those for ethyl propanoate (-0.53) and ethyl heptanoate (-1.37). It was tentatively suggested that for these two odorants the experimental psychometric functions may be in error, especially as the experimental psychometric function for butyl acetate does not seem to be compatible with our previously determined odour detection threshold (not determined for ethyl butanoate).

Table 15.4
C and D as obtained from the psychometric curves, and the fitted (calculated) values of C and D from Eqs. (15.4) and (15.5). The average absolute error (AAE) has also been calculated for both parameters

Compound	Method	C(psy)	D(psy)	C (eq. 4)	D (eq. 5)	AAE (C)	AAE (D)
ethanol	VDD	-0.7514	0.1948	-0.2478	0.2100	0.5036	0.0153
butan-1-ol	SB	-0.4467	0.4103	-1.0848	0.4412	0.6380	0.0309
butan-1-ol	VDD	-1.4680	0.2473	-1.0848	0.2072	0.3833	0.0402
heptan-2-one	SB	-1.1330	0.4504	-2.0455	0.5050	0.9125	0.0546
hexanoic acid	VDD	-2.8295	0.1492	-1.7848	0.1487	1.0448	0.0006
glutaraldehyde	VDD	-3.5599	0.4108	-2.8075	0.2989	0.7524	0.1119
butyl acetate	SB	-3.9939	0.4496	-1.5573	0.4982	2.4367*	0.0486
toluene	SB	-1.0088	0.5666	-0.5908	0.5671	0.4181	0.0006
ethyl propanoate	GV	-0.5330	0.4793	-1.1478	0.5003	0.6147	0.0210
ethyl heptanoate	GV	-1.3740	0.6031	-2.5923	0.4925	1.2183	0.1106
ethyl butanoate	VDD	-4.9054	0.5042	-1.4958	0.2638	3.4097*	0.2404*
benzene	VDD	0.1570	0.2779	-0.1865	0.3343	0.3435	0.0564
D-limonene	VDD	-1.3231	0.3130	-1.8543	0.3184	0.5312	0.0054
chloropicrin	VDD	-0.3198	0.2117	-0.3810	0.3149	0.0612	0.1033
naphthalene	GV	-3.1412	0.7044	-2.1508	0.6632	0.9905	0.0412
1-methyl naphthalene	GV	-2.7850	0.5810	-2.6925	0.6628	0.0925	0.0818
2-methyl naphthalene	GV	-2.7711	0.7458	-2.6453	0.6585	0.1259	0.0873
ozone	VDD	-1.5394	0.2691	2.3008	0.2363	3.8402*	0.0329
						0.5753	**0.0495**

* Not used in the calculations

The descriptors for TXIB (2,2,4-Trimethyl-1,3-pentanediol diisobutyrate), see Figure 15.6, that best fit the small number of experimental data that can be found for this compound have also been deduced: **E** = -0.047, **S** = 0.89, **A** = 0.00, **B** = 0.80, **L** = 8.842. From these descriptors C and D can be calculated, with C = -5.00 and D = 0.48 from Eq. (15.4) and Eq. (15.5), and then the psychometric function can be calculated. As shown in Figure 15.7, the calculated and experimental functions differ very greatly. At the moment, therefore, any prediction of psychometric functions is restricted to odorants whose calculated C and D values are normal as regards Figure 15.5.

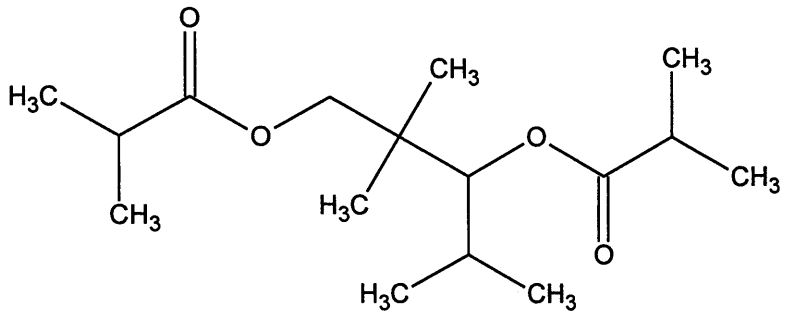


Figure 15.6. Structure of TXIB.

The C and D parameters for ethanol (GV) have also been calculated from Eq. (15.4) and Eq. (15.5), using the known descriptors, and it is found that C = -0.25 and D = 0.44. With these calculated C and D values the psychometric function can be constructed, as shown in Figure 15.7 where a comparison between the constructed function and experimental functions is made. Agreement is excellent. Hence the psychometric function for ethanol (GV) has been predicted just from the descriptors for ethanol, without any input at all from experimental odour data. Of course, it is expected that the calculated C and D values for ethanol (GV) will not be too far from the experimental C and D values for ethanol (VDD) that were used to obtain Eq. (15.4) and Eq. (15.5), and so this is not a very rigorous prediction. It does, however, serve as an example of how a psychometric function for an odorant can be predicted just starting with the physicochemical descriptors required in Eq. (15.4) and Eq. (15.5).

It is certainly necessary to measure detectability functions for an additional 3 or more dozen odorants (quite a time-consuming and labor intensive task) to have a comparable data set to that for odour thresholds, where the number of compounds is between 50 and 60 [18]. Nevertheless, here it is illustrated for more than a dozen odorants how Eq. (15.3) can be used to describe the constants C and D, which define a

complete detectability function, a much more comprehensive psychophysical outcome than a single threshold value.

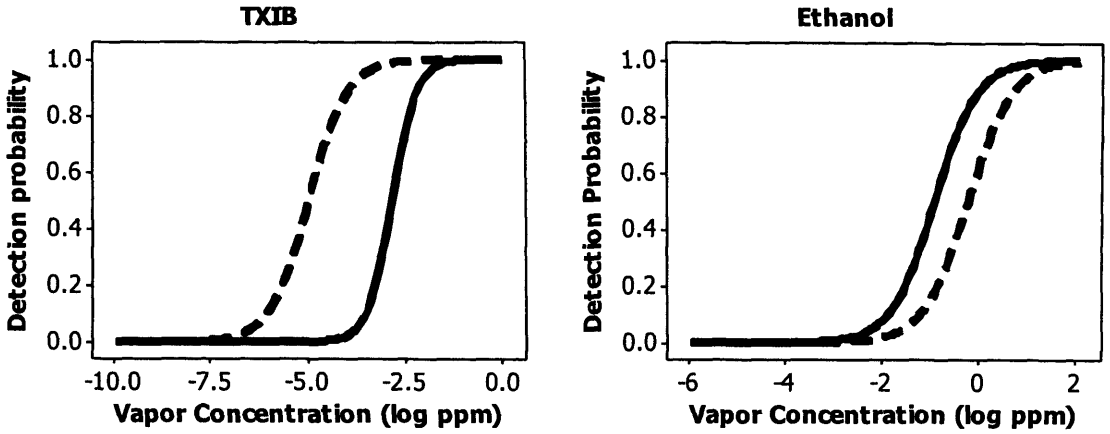


Figure 15.7. Odour psychometric functions for TXIB and ethanol, experimental (——) and calculated (– – – – –).

15.4. REFERENCES

1. K. Touhara, Microsc. Res. Tech. 58 (2002) 135
2. H. Breer, Biochem. Soc. Trans., 31 (2003) 113
3. Y. Niimura, M. Nei, Proc. Natl. Acad. Sci. U S A 100 (2003) 12235
4. T. Komiyama, L. Luo, Curr. Opin. Neurobiol. 16 (2006) 67
5. T. A. Schoenfeld, T. A. Cleland, Trends Neurosci. 28 (2005) 620
6. G.V. Ronnett, C. Moon, Annu. Rev. Physiol. 64 (2002) 189
7. W. B. Floriano, N. Vaidehi, W. A. Goddard, Chem Senses 29 (2004) 269
8. S. E. Hall, Floriano WB, Vaidehi N, and Goddard WA, Chem Senses 29 (2004) 595
9. K. Kajiya, K. Inaki, M. Tanaka, T. Haga, H. Kataoka, K. Touhara, J. Neurosci. 21 (2001) 6018
10. S. Katada, T. Hirokawa, Y. Oka, M. Suwa, K. Touhara, J. Neurosci. 25 (2005) 1806
11. M. S. Singer, Chem. Senses 25 (2000) 155
12. A. Kaneko, Keio J. Med., 50 (2001) 13
13. T . Abaffy, H. Matsunami, C. W. Luetje, J. Neurochem. 97 (2006) 1506
14. V. Matarazzo, O. Clot-Faybesse, B. Marcet, G. Guiraudie-Capraz, B. Atanasova, G. Devauchelle, M. Cerutti, P. Etievant, C. Ronin, Chem. Senses 30 (2005) 195
15. M. Chastrette, SAR QSAR Environ. Res. 6 (1997) 215
16. K. J. Rossiter, Chem. Rev. 96 (1996) 3201

17. T. Yamanaka, Chem. Senses 20 (1995) 471
18. M. H. Abraham, J. M. Ǵola, J. E. Cometto-Muniz, W. S. Cain, Chem Senses 27 (2002) 95
19. K. M. Hau, D. W. Connell, Indoor Air 8 (1998) 23
20. Personal communication from Enrique Cometto-Muñiz
21. W. S. Cain, Ear Nose Throat J., 68 (1989) 316
22. W. S. Cain, J. E. Cometto-Muñiz, R. A. de Wijk (1992) *Techniques in the quantitative study of human olfaction.* In: M. J. Serby, K. L. Chobor, ed. Science of Olfaction. New York: Springer-Verlag, 279-308
23. R. Schmidt, W.S. Cain, Chem. Senses, 28 (2003) 560 (abstract)
24. J. E. Cometto-Muñiz, W. S. Cain, T. Hiraishi, M. H. Abraham, J. M. R. Gola, Chem. Senses 25 (2000) 285
25. H. N. Knudsen, G. Clausen, P. O. Fanger, Indoor Air 7 (1997) 107
26. H. N. Knudsen, O. Valbjørn, P. A. Nielsen, Indoor Air 8 (1998) 264
27. D. G. Laing, Perception 11 (1982) 221
28. D. G. Laing, Perception, 12 (1983) 99
29. N. A. Macmillan, C. D. Creelman (1991) *Detection theory: A user's guide.* Cambridge: Cambridge University Press
30. M. Devos, F. Patte, J. Rouault, P. Laffort, L. J. van Gemert (1990) Standardized Human Olfactory Thresholds. Oxford: IRL Press
31. J. E. Cometto-Muñiz (2001) *Physicochemical basis for odor and irritation potency of VOCs.* In: J. D. Spengler, J. Samet, J. F. McCarthy, ed. Indoor Air Quality Handbook. New York: McGraw-Hill, 20.21-20.21
32. J. E. Cometto-Muñiz, W. S. Cain, Arch. Environ. Health 48 (1993) 309
33. M. H. Abraham, Chem. Soc. Rev. 22 (1993) 73
34. M. H. Abraham, P.K. Weathersby, J. Pharm. Sci. 83 (1994) 1450
35. M. H. Abraham, J. M. Gola, J. E. Cometto-Muniz, W. S. Cain, Indoor Built Environ 10 (2001) 252
36. M. H. Abraham, M. Hassanisadi, M. Jalali-Heravi, T. Ghafourian, W.S. Cain, J. E. Cometto-Muñiz, Toxicol. Sci. 76 (2003) 384
37. M. H. Abraham, R. Kumarsingh, J. E. Cometto-Muñiz, W. S. Cain, Arch. Toxicol. 72 (1998) 227
38. M. H. Abraham, J. Andonian-Haftvan, G. S. Whiting, A. Leo, R. S. Taft, J. Chem. Soc. Perkin Trans. 2 (1994) 1777

39. M. H. Abraham, J. Le, W. E. Acree Collect Czech Chem Commun, 64 (1999) 1748
40. M. H. Abraham, J. A. Platts, A. Hersey, A. J. Leo, R. W. Taft, J. Pharm. Sci. 88 (1999): 670
41. M. H. Abraham, G. S. Whiting, P. W. Carr, H. Ouyang, J. Chem. Soc. Perkin Trans 2 (1998) 1385
42. M. H. Abraham, C. F. Poole, S. K. Poole, J. Chromatogr. A 842 (1999) 79

Chapter 16 Concentration-Detection Functions for Nasal Pungency Detectability of Airborne Chemicals

16.0. INTRODUCTION

The nasal cavity is innervated by two major brunches of the trigeminal nerve: the ophthalmic branch and the maxillary branch [1], which convey purely sensory information. Sensations derived from the trigeminal nerve are somatosensory, i.e. touch, temperature or pain sensations and perception of atmospheric humidity [2]. These feelings are, as mentioned in Chapter 2, stinging, piquancy, burning, tingling, freshness, prickling, irritation, itching, cooling. Two major fiber systems, the unmyelinated C-fibers and the myelinated Aδ-fibers, participate in the afferent chemosensitive innervation of the nasal respiratory epithelium [3]. In sensory nerve, the signal transduction mechanisms underlying perception of chemical stimuli are not fully understood. Although several receptors have been identified, this question remains widely uncovered. According to microneurographic studies, the 'burning' sensation is related to the activation of C-fibers and the 'stinging' sensation to the activation of Aδ-fibers [4]. Both fibers are activated by the intracellular accumulation of protons which modify the membrane conductance [5], by increasing cation membrane conductance [6,7] which generally exhibits slow desensitization [8].

In the present study concentration-detection, i.e., detectability, functions for the human olfactory detection of six chemical vapours of diverse chemical structures measured by Cometto-Muñiz and co-workers in the past few years have been gathered [9]. The functions cover the complete peri-threshold range from chance to almost perfect detection and provide much more information than what a single "threshold value" can convey. In this chapter, these nasal pungency functions have been modeled with the same uniform approach used in Chapter 15, and have been correlated with the same set of up to five physicochemical descriptors taken from the general solvation equation of mechanistic significance that has quite successfully described the simpler "nasal pungency threshold" psychophysical outcome, as described in Chapter 12 of this work.

16.1. MATERIALS AND METHODS

16.1.1 Subjects

All participants studied by Cometto-Muñiz and co-workers were anosmic (i.e., lacked a sense of smell) as determined by a clinical olfactory test [10]. Table16.1 describes the characteristics of all subjects tested. In this table the chemicals tested, the delivery techniques, the number of subjects (# S's), their average age (±SD) and range, the gender distribution (Females(F)/Males(M)), the number of trials per subject/concentration, and the total number of trials per concentration, can also be found. Delivery techniques include: squeeze bottle (SB) and glass vessel (GV), see below [9].

Table 16.1
Characteristics of all subjects tested [9]

Chemical Stimulus	Delivery Technique	# S's	Average Age(±SD)	Age Range	F/M	Trials per Subject	Total trials
1. 1-butanol	SB	4	40(±14)	28-59	3/1	16	64
2. 2-heptanone	SB	4	40(±14)	28-59	3/1	16	64
3. butyl acetate	SB	4	44(±13)	29-60	3/1	16	64
3. butyl acetate	GV	5	51(±15)	34-71	3/2	≥19	96
4. ethyl propanoate	GV	5	44(±20)	20-64	2/3	20	100
5. ethyl heptanoate	GV	5	44(±20)	20-64	2/3	20	100
6. toluene	SB	4	44(±13)	29-60	3/1	16	64
6. toluene	GV	5	51(±15)	34-71	3/2	≥19	96

16.1.2 Stimuli and Equipment

Six stimuli were tested (Table 16.1). All chemicals were high purity, typically >99%, as provided by the chemical suppliers. Whenever available, those meeting Food Chemical Codex (FCC) quality were chosen. In most cases, their delivered vapour concentrations were confirmed analytically by gas chromatography (GC), high performance liquid chromatography (HPLC) or a chemical-specific instrument (e.g., ozone analyzer). In a few cases, concentrations were calculated from total mass of chemical evaporated and

volume of dilution air or nitrogen. All concentrations are expressed as log ppm by volume. Presentation of stimuli [11] involved a static system via squeeze bottles (SB) and/or glass vessels (GV) [12] (see Table 16.1). For odour stimulation with a static system, SB and GV ended, respectively, in a single spout or two nosepieces from which the subjects sniffed the headspace of the container [13]. (When using SB, subjects were instructed to squeeze with approximately equal strength on all trials.)

16.1.3 Procedure

All odour testing involved using a two- or three-alternative forced-choice procedure between stimulus and blanks [14]. Blanks comprised the headspace above mineral oil (light, FCC) carried by either nitrogen or air.

16.1.4 Data Analysis

The outcome is summarized in terms of nasal pungency detection probability, i.e., detectability, as a function of stimulus vapour concentration (log ppm). Detectability was corrected for chance according to [15]:

$$P = [m \cdot p(c) - 1] / (m - 1) \tag{16.1}$$

where P = detection probability corrected for chance, m = number of choices in the forced-choice procedure (i.e., 2 or 3), and p(c) = proportion correct (i.e., number of correct trials / total number of trials).

In the present work, nasal pungency concentration-detection functions were modeled by the same simplified sigmoid equation used in the previous chapter, of the form:

$$y = 1 / (1 + e^{[-(x-C)/D]}) \tag{16.2}$$

where y = detectability (P) as defined in Eq. (16.1), x = vapor concentration of the chemical stimulus (in log ppm by volume), and C and D are constants. Note that C represents the concentration of the stimulus (i.e., x) when y = 0.5, that is, when detectability is half way (i.e., P =0.5) between chance detection (i.e., P = 0.0) and

perfect detection (i.e., P = 1.0). In turn, the value of constant D governs the steepness of the detectability function.

16.2. RESULTS

Figure 16.1 presents nasal pungency functions for the six chemicals. Table 16.2 lists the values for the constants C and D, with their respective standard errors, and the values for Chi-square and the coefficient of determination R^2 of the functions. Once again, both the figure and the table show that Eq. (16.2) provides an excellent description of the experimental data.

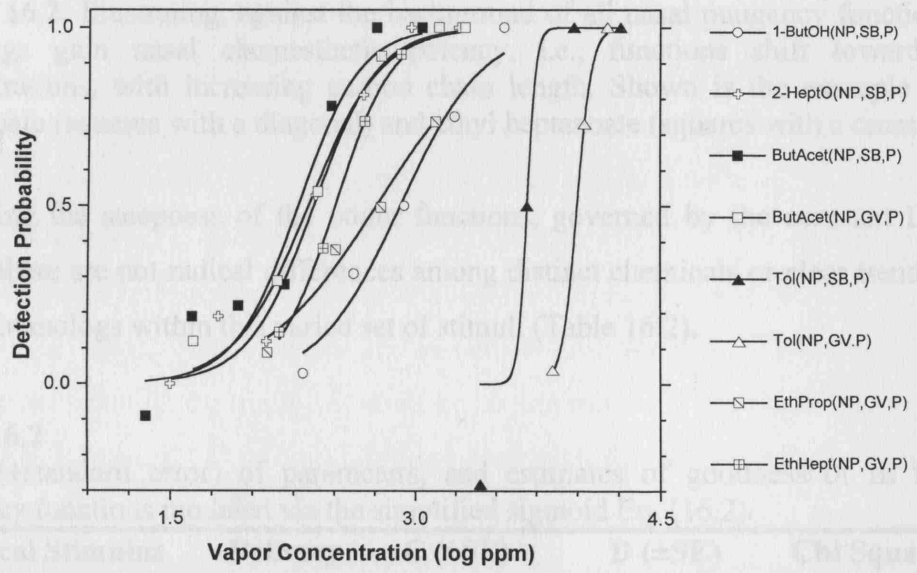


Figure 16.1. Psychometric functions for nasal pungency modeled by Eq. (16.2) for chemicals and techniques listed in Table 16.2.

As mentioned under Data analysis, the value of C represents the concentration of the chemical producing a level of detectability situated half way between chance detection and perfect detection. Such a concentration is commonly, although not always, taken as the threshold concentration. Again the testing of homologous chemical series was chosen to assess the relative nasal chemesthetic potency of vapours.

Odour potency across homologs tested under identical method and procedure (Table 16.2) can be compared for the following esters, both tested via GV (ethyl propanoate and ethyl heptanoate). The functions clearly show the expected trend: the larger homolog has a higher potency, as reflected by curves shifted to the left, i.e., towards lower concentrations (Figure 16.2).

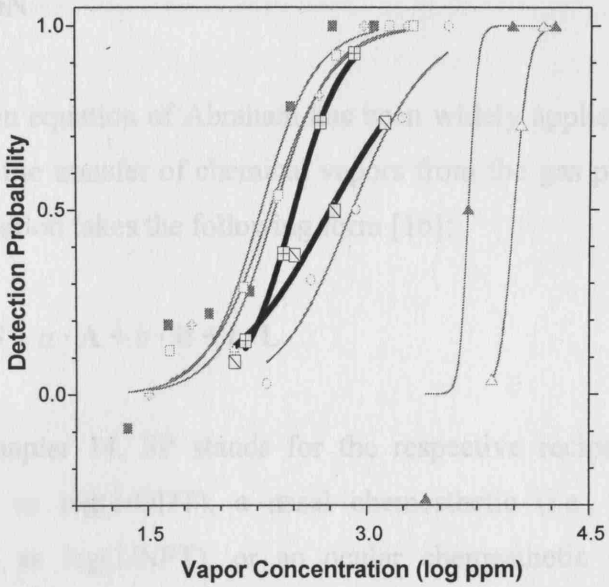


Figure 16.2. Illustrating, against the background of all nasal pungency functions, how homologs gain nasal chemesthetic potency, i.e., functions shift towards lower concentrations, with increasing carbon chain length. Shown is the example of ethyl propanoate (squares with a diagonal) and ethyl heptanoate (squares with a cross).

Regarding the steepness of the odour functions, governed by the constant D on Eq. (16.2), there are not radical differences among distinct chemicals or clear trends among related homologs within this varied set of stimuli (Table 16.2).

Table 16.2
Value (±standard error) of parameters, and estimates of goodness of fit for nasal pungency functions modeled via the simplified sigmoid Eq. (16.2)

Chemical Stimulus	Delivery Technique	C (±SE)	D (±SE)	Chi Square	R^2
1. butan-1-ol	SB	2.91 (±0.048)	0.25 (±0.045)	0.015	0.97
2. heptan-2-one	SB	2.37 (±0.043)	0.20 (±0.039)	0.027	0.97
3. butyl acetate	SB	2.28 (±0.063)	0.19 (±0.056)	0.063	0.95
3. butyl acetate	GV	2.34 (±0.030)	0.22 (±0.032)	0.0090	0.99
4. ethyl propanoate	GV	2.76 (±0.042)	0.35 (±0.050)	0.0046	0.98
5. ethyl heptanoate	GV	2.51 (±0.010)	0.17(±0.0096)	0.00080	1.00
6. toluene	SB	3.69 (±0.23)	0.026 (±17.6)	0.079	0.93
6. toluene	GV	4.00(±0.005)	0.044(±0.004)	0.00025	1.00

16.3. DISCUSSION

The general solvation equation of Abraham has been widely applied for the correlation of process entailing the transfer of chemical vapors from the gas phase to a condensed (bio)phase. This equation takes the following form [16]:

$$SP = c + e \cdot \mathbf{E} + s \cdot \mathbf{S} + a \cdot \mathbf{A} + b \cdot \mathbf{B} + l \cdot \mathbf{L} \qquad (16.3)$$

As described in Chapter 14, SP stands for the respective reciprocals of: an odour detection threshold as log(1/ODT), a nasal chemesthetic (i.e., nasal pungency or irritation) threshold as log(1/NPT), or an ocular chemesthetic (i.e., eye irritation) threshold as log(1/EIT), where all thresholds are expressed as log ppm by volume [17,18,19,20]. The bold letters in Eq. (16.3) represent physicochemical descriptors of volatile compounds as follows: **E** is an excess molar refraction, **S** is the dipolarity/polarizabiltiy, **A** and **B** are the overall or effective hydrogen-bond acidity and basicity, respectively, and **L** (log L^{16}) is defined through L^{16}, the volatile compound (i.e., the solute) Ostwald solubility coefficient on hexadecane at 298 K. The **L**-descriptor is itself a combination of two solute properties: (i) a general measure of solute size, and (ii) the ability of a solute to interact with a solvent phase (i.e., the receptor environment or biophase) through dispersion forces. The units of **E** are $cm^3/10$ whereas the other descriptors have no units because they are all derived from the logarithm of an equilibrium constant. In turn, the coefficients *c*, *e*, *s*, *a*, *b*, and *l* are found by multiple linear regression analysis, Nevertheless, they are not simply fitting coefficients, because they reflect the complementary properties that must characterize the receptor biophase for an optimum interaction with the solute, that is, the dissolved odorant. In this way, the solvation Eq. (16.3) conveys mechanistic information on the interaction of the odorant (or solute) with the odorant-receptor environment (or solvent) as follows: The *e*-coefficient gives the tendency of the receptor biophase to interact with the odorant through polarizability-type interactions, mostly via electron pairs. The *s*-coefficient is a measure of the receptor biophase dipolarity/polarizability. The *a*-coefficient represents the complementary property to the odorant hydrogen-bond acidity and thus is a measure of the receptor biophase hydrogen-bond basicity. Likewise, the *b*-coefficient is a measure of the receptor biophase hydrogen-bond acidity. Finally, the *l*-coefficient is a measure of the hydrophobicity of the receptor biophase.

There is not much that can be done with only six compounds and a data set of eight observations. The physicochemical descriptors of the corresponding compounds can be found in Table 16.3. The equations for C and D of the nasal pungency psychometric functions are as follows:

$$C = 2.27 + 2.61\ \mathbf{E} \qquad (16.4)$$

$N = 8$; $SD = 0.199$; $R^2 = 0.920$; $F = 68.97$

$$D = 0.839 + 2.45\ \mathbf{E} - 6.05\ \mathbf{S} - 4.29\ \mathbf{A} + 6.24\ \mathbf{B} \qquad (16.5)$$

$N = 8$; $SD = 0.025$; $R^2 = 0.976$; $F = 29.98$

The statistics of the equations seem to be good enough to use them to predict further values of C and D, equations (16.4) and (16.5) respectively. The L descriptor is not included in any of the equations and this is suspicious because the cavity term should be important in the process of the compound reaching the receptor site. Probably, in a larger data set this term will appear in the equations.

Table 16.3
Descriptors for the set of compounds used to build Eqs. (16.4) and (16.5)

Chemical Stimulus	E	S	A	B	L
butan-1-ol	0.224	0.42	0.37	0.48	2.601
2-heptanone	0.123	0.68	0.00	0.51	3.760
butyl acetate	0.071	0.60	0.00	0.45	3.353
ethyl propanoate	0.087	0.58	0.00	0.45	2.807
ethyl heptanoate	0.027	0.58	0.00	0.45	4.733
toluene	0.601	0.52	0.00	0.14	3.325

Using the same range of vapor concentration, in log ppm unit, used in the experiments, the C and D parameters have been calculated, using equations (16.4) and (16.5), for the same compounds for which the C and D parameters were obtained experimentally, with the aim of comparing the shape of the experimental and predicted graphs, and see whether the equations will allow the prediction of C and D.

As shown in Figure 16.3, it seems that equations (16.4) and (16.5) can be used to predict C and D respectively. The curves are very close one to each other and the shapes are very similar. For butan-1-ol, the experimental and calculated functions even overlap.

The plots in Figure 16.3 suggest that it is possible to predict the entire psychometric functions for an irritant simply by predicting the relevant C and D values from Eqs. (16.4) and (16.5), using already known values of the descriptors.

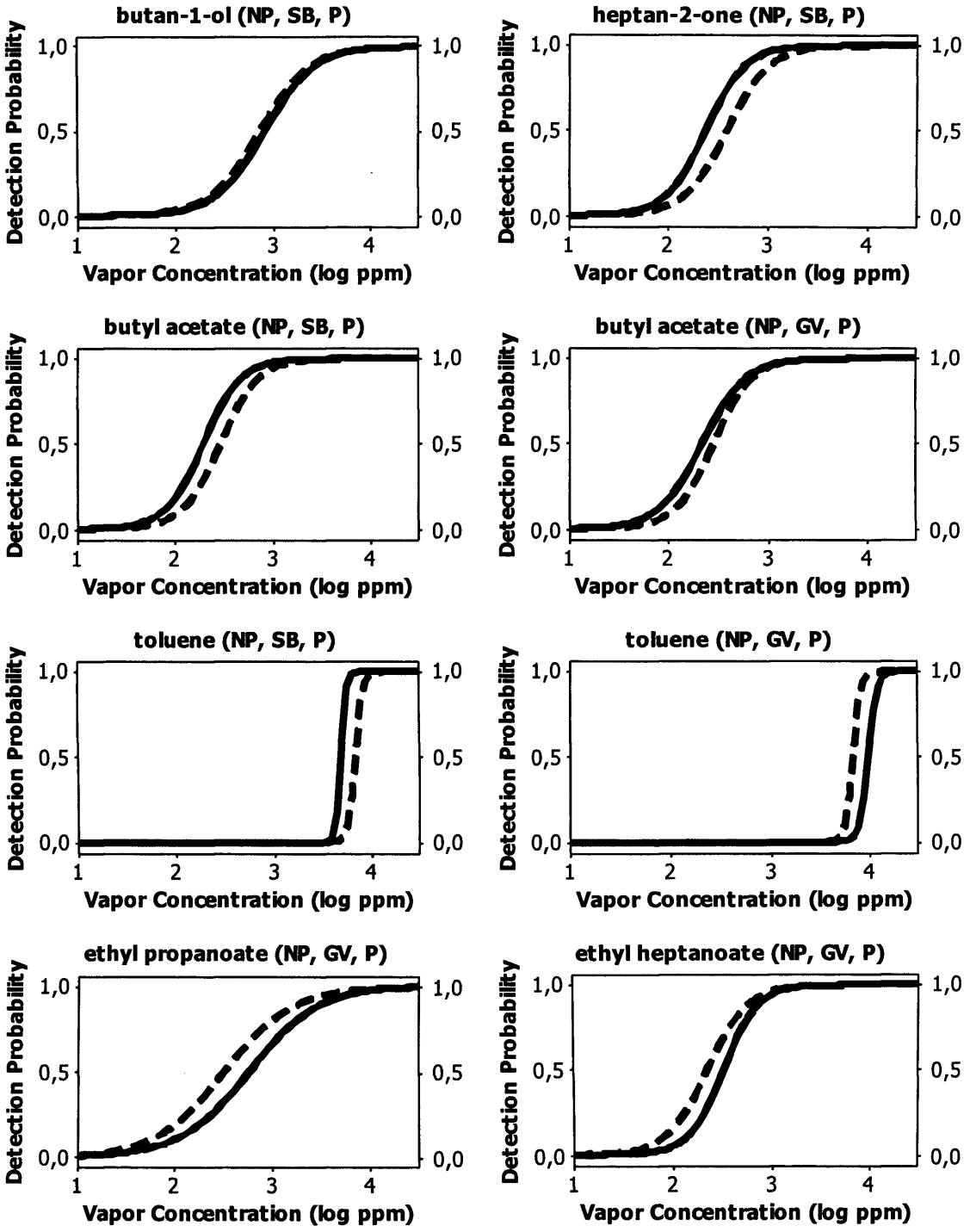


Figure 16.3. Comparison of the experimental (————) and calculated (– – – – –) nasal pungency psychometric functions.

The experimental values of C and D are plotted in Figure 16.4. A peculiar case is that of toluene whose C and D values lie far from the rest of the compounds in both experiments.

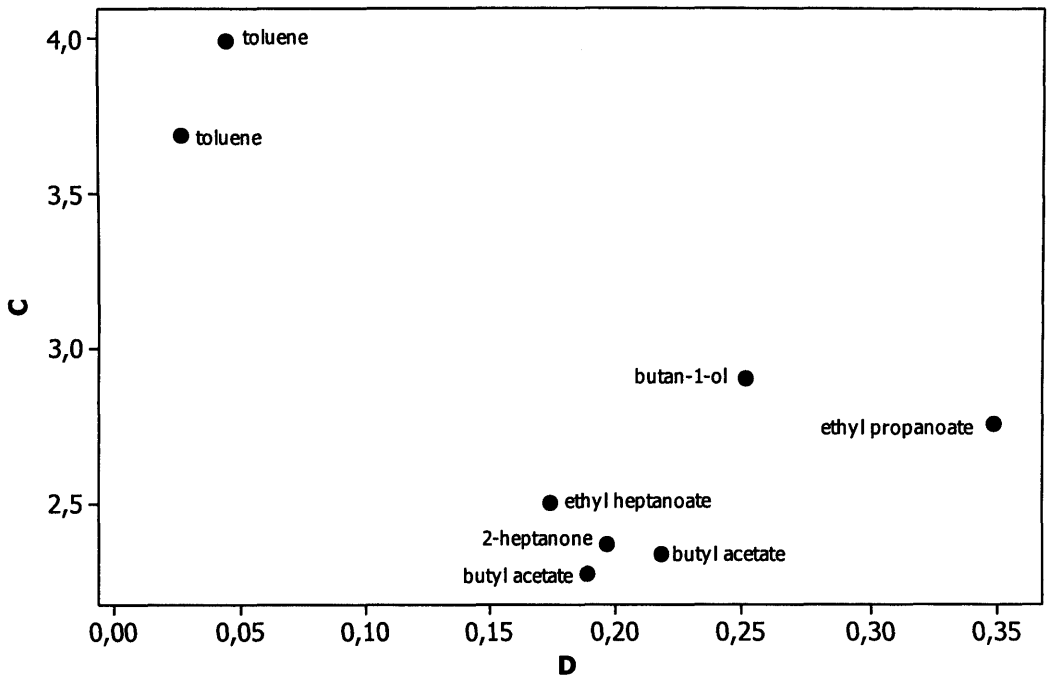


Figure 16.4. Representation of the experimental values of C against D.

However, the small number of compounds in the data set, and the fact that the predicted C and D values come both from two different methods (i.e., SB and GV), see Table 16.2, makes it difficult to decide whether there was an error in the experiments with this compound or it is just that the range of the constants was not covered at all.

16.4. REFERENCES

1. J. Lang, (1989) Clinical Anatomy of the Nose, Nasal Cavity and Paranasal Sinuses. Thieme Verlag, Stuttgart.
2. J.P. Kelly, J. Dodd, (1991) *Trigeminal system.* In Kandel,E.R., Schwartz,J.H. and Jessell,T.M. (eds.), Principles of Neural Science. Elsevier, New York, pp. 701-710.
3. S. I. Sekizawa, H. Tsubone, Resp. Physiol. 96 (1994) 37
4. T. Hummel, A. Livermore, C. Hummel, G. Kobal, Electroen. Clin. Neuro. 84 (1992) 192

5. K. H. Steen, H. Wegner, H. Kreysel, P. Reeh, Society of Neuroscience 21 (1995) 648.
6. S. Bevan, J.C. Yeats, J. Physiol.-London 433 (1991) 145
7. S. Bevan, C.A. Forbes, Winter J, J. Physiol.-London 459 (1993) P401
8. K.H. Steen, P.W. Reeh, F. Anton, H.O. Handwerker, J. Neurosci. 12 (1992) 86
9. Personal communication from Enrique Cometto-Muñiz
10. W.S. Cain, Ear Nose Throat J., 68 (1989) 316
11. W. S. Cain, J. E. Cometto-Muñiz, R. A. de Wijk (1992) *Techniques in the quantitative study of human olfaction.* In: M. J. Serby, K. L. Chobor, ed. Science of Olfaction. New York: Springer-Verlag, 279-308
12. R. Schmidt, W.S. Cain, Chem. Senses 28 (2003) 560 (abstract)
13. J. E. Cometto-Muñiz, W. S. Cain, T. Hiraishi, M. H. Abraham, J. M. R. Gola, Chem. Senses 25 (2000) 285
14. J. E. Cometto-Muñiz, W. S. Cain, M. H. Abraham, J. M. R. Gola, J. Appl. Toxicol. 22 (2002) 25
15. N. A. Macmillan, C. D. Creelman, (1991) Detection theory: A user's guide. Cambridge University Press, Cambridge.
16. M. H. Abraham, Chem. Soc. Rev. 22 (1993) 73
17. M. H. Abraham, J. M. Gola, J. E. Cometto-Muñiz, W. S. Cain, Chem. Senses 27 (2002) 95
18. M. H. Abraham, J. M. Gola, J. E. Cometto-Muñiz, W. S. Cain, Indoor Built Environ. 10 (2001) 252
19. M. H. Abraham, M. Hassanisadi, M. Jalali-Heravi, T. Ghafourian, J. E. Cometto-Muñiz, W. S. Cain, Toxicol. Sci. 76 (2003) 384
20. M. H. Abraham, R. Kumarsingh, J. E. Cometto-Muñiz, W. S. Cain, Arch. Toxicol. 72 (1998) 227

Chapter 17 Chemical boundaries for detection of eye irritation in humans from homologous vapors

17.1. INTRODUCTION

Previous studies of ocular and nasal chemesthetic thresholds along and across homologous chemical series have suggested the existence of a "cut-off", i.e., a point where a homolog fails to stimulate even at vapor saturation [1]. All homologs larger than the 'cut-off' compound fail as well. The phenomenon has been observed for other biological effects of VOCs, for example, anesthesia [2]. The cited QSAR in Chapter X.X for chemesthesis would predict a cut-off only if the calculated threshold came out higher than the saturated vapor concentration of the homolog at room temperature. In this case, one could heat the liquid or solid VOC to increase its vapor concentration allowing it to reach chemesthetic detection. Experiments on ocular and nasal chemesthesis have applied this strategy and found that most VOCs continue to fail to reach detection [3,4]. The outcome suggests that, in these cases, the cut-off is not concentration-based but rather molecular dimension- or structure-based. For example, the VOC could be too large or bulky to interact or fit into the receptor pocket of a protein. The present study sought, first, to identify the particular cut-off member along homologous alkylbenzenes and 2-ketones and, second, to probe the likely basis for the effect. The following experiments, Experiment 1 and Experiment 2, were carried out by Enrique Cometto-Muñiz and co-workers.

17.2. EXPERIMENT 1: DETECTION OF EYE IRRITATION AT ROOM TEMPERATURE (23°C)

17.2.1 Materials and Methods

17.2.1.1 Subjects

Fifteen participants (9 females) took part in the experiment. Their average age (±SD) equaled 21 (±4) years and ranged from 18 to 35 years. Thirteen of them (8 females) performed in the normosmic range on a clinical olfactory test [5]. The other two performed in the very mildly hyposmic range for either one nostril (a male) or for both (a female). All participants were non-smokers. Thirteen subjects did not use contact

lenses. Of the remaining two, one (female) used contact lenses regularly and another (male) used them occasionally; neither subject wore contact lenses on testing days.

17.2.1.2. Stimuli and Equipment

Based on previous research on nasal chemesthesis [3] and on preliminary testing, two members from a 2-ketone homologous series and three from an alkylbenzene series were selected. They were (purity in parenthesis): 2-dodecanone (98+%) and 2-tridecanone (≥99.6%), and pentyl (or amyl) benzene (98%), hexyl benzene (98%) and heptyl benzene (98%), respectively. These compounds were found to approach a "cut-off" point within each series such that even vapor from the neat homolog failed to evoke obvious eye irritation. Mineral oil (light, Food Chemical Codex quality) served as blank. Chemical vapors and blanks were delivered from a 1,900 ml glass vessel system containing 200 ml of neat stimuli (50 g for the solid 2-tridecanone) or blank and ending in a bowl-shaped 40-ml eyepiece, as described previously [4]. To avoid depletion of the headspace in the vessels, we prepared each chemical in duplicate, alternating their use through the session. The headspace concentration in a vessel was measured by gas chromatography (flame ionization detector) via a calibration curve for mass specific for each chemical [6]. The resulting concentrations in ppm by volume (±SD) were: 19.5 (±1.1) for 2-dodecanone, 5.5 (±0.53) for 2-tridecanone, 355 (±12) for pentyl benzene, 128 (±28) for hexyl benzene, and 27.5 (±1.4) for heptyl benzene.

17.2.1.3. Procedure

A trial consisted of a three-alternative forced-choice (3AFC) against blanks, where subjects had to decide which ocular exposure from a triad produced a different, typically stronger, sensation. The position of the stimulus within the triad and the order of testing of the five chemicals were randomized. Time of exposure was constant at 6 sec for all presentations. Flowrate to the eye was set to either 4 L/min or 8 L/min, in irregular order, alternating 20 trials at 4 L/min and 20 at 8 L/min. Thus, there were 10 different stimuli (5 chemicals × 2 flowrates). Participants wore noseclips during trials to avoid odor cues, and alternated the exposed eye (right or left) from one triad to the next. An interval of at least one minute elapsed between such successive triads testing different eyes. The vessels in a triad were presented covered by an opaque plastic sleeve to avoid

visual clues. Subjects were instructed to end the exposure in the unlikely event that clear irritation was felt, and, importantly, not to proceed with the next member of a triad until all previous sensations (if any) had disappeared completely. We stress that stimuli were specifically chosen to be at the very border of chemesthetic detection/no detection and that these low and very brief levels of stimulation precede any clinical ocular signs [7]. In addition, participants rated their confidence on the decision made for each triad, using a scale ranging from "1" (not confident at all) to "5" (extremely confident). Subjects participated in 4 to 6 sessions, each lasting between 2 and 3 hours, until they accumulated 20 to 26 judgments per stimulus per subject. This provided a group total of 312 to 316 group judgments per stimulus.

17.2.1.4. Data Analysis

Results are presented as detection probability (i.e., detectability) or as confidence of detection as a function of the stimuli. Detection probability was corrected for chance using the following formula [8]:

$$P = (m \cdot p(c) - 1) / (m - 1) \tag{17.1}$$

where P = detectability corrected for chance, m = number of choices per trial (in our case, three), and p(c) = proportion correct (i.e., number of correct trials / total number of trials). Statistical significance was established by analysis of variance (ANOVA) for repeated measurements (Software: SuperANOVA v.1.11, Abacus Concepts, Inc., Berkeley, CA).

17.2.2 Results

Figure 17.1 summarizes the results for the group in terms of detectability and confidence of detection for each stimulus. As observed along other homologous series [4], chemesthetic detectability decreased with increasing carbon chain length, and confidence of detection followed this trend closely. Increasing the flowrate to the eye produced a slight increase in detectability but only for vapors detectable near the middle range of the scale (i.e., $0.35 \leq P \leq 0.65$), for example, hexyl benzene and 2-dodecanone,

not for those that were very detectable (i.e., $P \geq 0.75$), for example, pentyl benzene, or for those that were practically undetectable (i.e., $P \leq 0.15$), for example, heptyl benzene and 2-tridecanone. In other words, an increased flowrate did not modify the detectability of readily detectable stimuli and, very importantly, not that of almost undetectable stimuli.

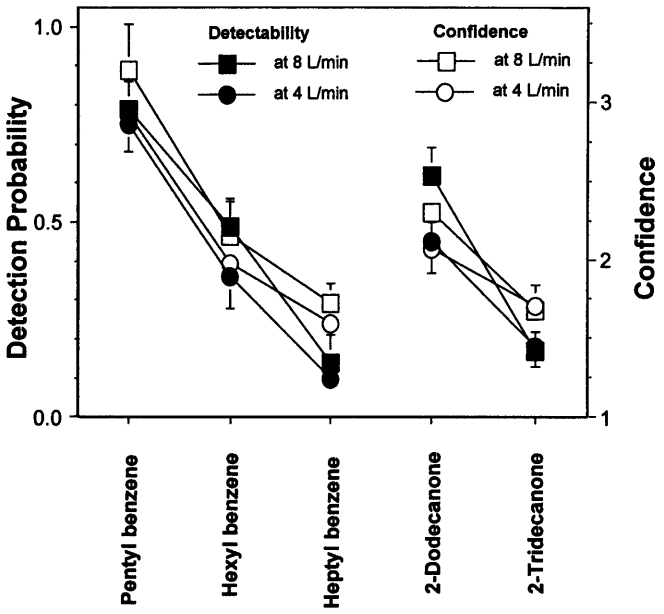


Figure 17.1. Detectability (left *y*-axis) and confidence of detection (right *y*-axis) of the group (n=15) for the five chemicals presented to the eye at two flowrates. Each point represents the average of 312 to 316 judgments provided by 15 subjects. Bars indicate standard error (SE).

In an ANOVA for repeated measurements with the factors flow (2 levels: 4 and 8 L/min) and chemical (5 levels), the factor flow was significant ($F(1,14) = 8.376$, $p = 0.0118$) and so was the factor chemical ($F(4,56) < 0.0001$), but not their interaction. A look at the individual data reveals that, with minor variations, all subjects showed the trend observed for the group (Figure 17.2).

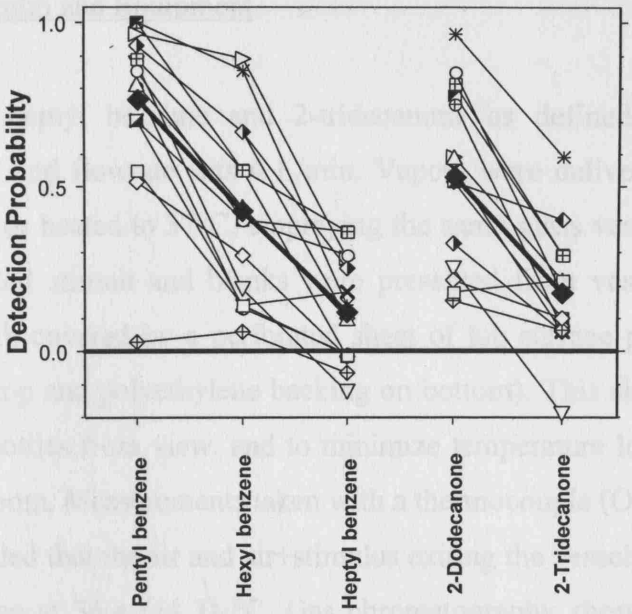


Figure 17.2. Individual ocular detectability of each chemical, averaged across the two flowrates of presentation, for the 15 subjects (dotted lines, each symbol represents one subject). Single data points comprise between 40 and 54 judgments per chemical. The average data for the group is also shown (thick, continuous lines, filled diamonds). Group data points comprise between 624 and 630 judgments per chemical. The thick horizontal line highlights chance detection level.

17.3. EXPERIMENT 2: COMPARISON OF THE EYE IRRITATION DETECTABILITY OF HEPTYL BENZENE AND 2-TRIDECANONE AT 23 AND AT 37 °C

17.3.1 Materials and Methods

17.3.1.1. Subjects

Twenty-one participants (11 females) took part in the experiment. Their average age (±SD) equaled 29 (±10) years and ranged from 19 to 51 years. Seventeen of them (10 females) performed in the normosmic range on a clinical olfactory test [5]. The other four performed in the very mildly hyposmic range for either one nostril (2 males) or for both (1 female, 1 male). All participants but one (a male) were non-smokers. Seventeen subjects did not use contact lenses. Of the remaining four, one (female) used them regularly, another (male) used them occasionally, and the other two (males) were former users. No subject wore contact lenses on testing days. Three subjects (2 males, 1 female) had also participated in Experiment 1.

17.3.1.2. Stimuli and Equipment

Stimuli comprised heptyl benzene and 2-tridecanone as defined before. Time of exposure was 6 sec and flowrate was 4 L/min. Vapors were delivered either at 23°C (room temperature) or heated to 37°C, employing the same glass vessel system used in Experiment 1. Heated stimuli and blanks were presented from vessels standing in a calibrated water bath covered by a perforated sheet of lab surface protector (made of cellulose fibers on top and polyethylene backing on bottom). This sheet served to hide the content of the bottles from view, and to minimize temperature loss from the water bath to the testing room. Measurements taken with a thermocouple (Omega Instruments, Stanford, CT) revealed that the air and air+stimulus exiting the vessels in our conditions of stimulation did so at 36.4 (±1.3) °C. Gas chromatography showed that the vapor concentration (ppm by volume ±SD) of heated heptyl benzene was 141 (±15) and that of heated 2-tridecanone was 38 (±3). These values are 5.0 and 6.9 times higher, respectively, than the corresponding vapor concentrations at 23°C (see Experiment 1).

17.3.1.3. Procedure

Analogous to that in Experiment 1 (i.e., 3AFC and rating of confidence) except that in Experiment 2 there were only 4 different stimuli (2 chemicals X 2 temperatures), as described above. Subjects participated in 2 to 3 sessions, each lasting between 2 and 3 hours, until 20 judgments per stimulus per subject were gathered. This provided a group total of 420 judgments per stimulus.

17.3.1.4. Data Analysis

Same as in Experiment 1.

17.3.2 Results

Figure 17.3 summarizes the outcome of Experiment 2 and puts it in context with that of Experiment 1. The graph shows agreement between both experiments in terms of the detectability of heptyl benzene and 2-tridecanone at room temperature, 23°C. It also shows that heating the VOCs to 37°C barely increased their detectability. An analysis of

variance (ANOVA) for repeated measurements on the data from Experiment 2 including the factors chemical (two levels: heptyl benzene and 2-tridecanone) and temperature (two levels: 23 and 37 °C) revealed that neither factor nor their interaction was statistically significant ($p>0.10$). We conclude that, in terms of group data, the increased vapor concentration of both VOCs heated to 37°C failed to increase their detectability by eye irritation.

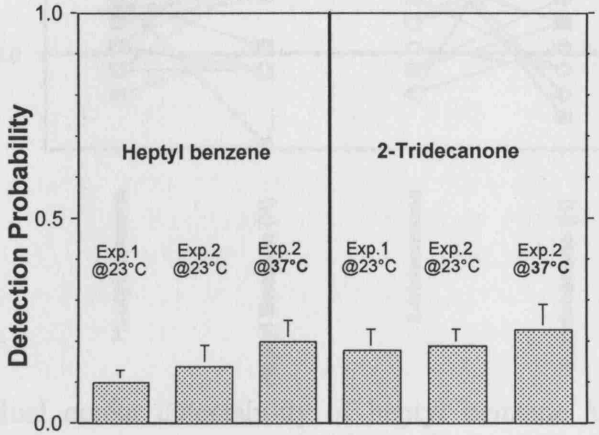


Figure 17.3. Showing: a) agreement in the ocular detectability of heptyl benzene and 2-tridecanone between the group in Experiment 1 (n=15) and that in Experiment 2 (n=21), both tested at room temperature, 23°C, and b) failure to enhance significantly the detectability of either chemical by increasing their vapor concentration via heating them to 37°C.

A more detailed level of understanding can be achieved by looking at the individual data (Figure 17.4). Based on the individual variability observed in Figure 17.4, the analysis that follows will consider that: 1) values of detectability (i.e., P) below P=0.20 imply virtually no detection, and 2) only increases in detectability of P>0.20 and falling above P=0.20 constitute relevant increases. Under these terms, and in contrast with the uniform trend for individual data seen in Figure 17.2, we can observe in Figure 17.4 two types of individual trends for each chemical: 1) for most subjects (13 out of 21 for both heptyl benzene and 2-tridecanone) the increase in concentration produced by heating to 37°C failed to either precipitate or enhance detection, and 2) for the rest of the subjects (8 out of 21 for both chemicals) the increase in concentration did precipitate detection or enhanced it. Most of the 13 individuals for whom increased concentration did not enhance detection were common for both chemicals (9 out of the 13). Many of the 8

individuals for whom concentration did enhance detection were common for both chemicals (4 out of the 8).

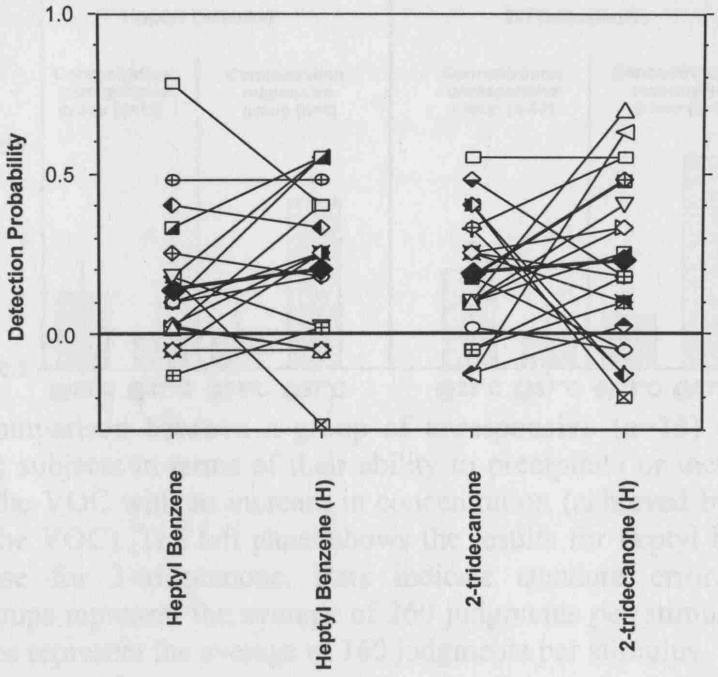


Figure 17.4. Individual ocular detectability of heptyl benzene and 2-tridecanone at room temperature, 23°C, and heated (H) to 37°C, for the 21 subjects (dotted lines, each symbol represents one subject). Single data points comprise 20 judgments per stimulus. The average data for the group is also shown (thick, continuous lines, filled diamonds). Group data points comprise 420 judgments per stimulus. The thick horizontal line highlights chance detection level.

Figure 17.5 illustrates, for each chemical, a comparison in ocular chemesthetic detectability between the group unresponsive to an increase in concentration (n=13) and the group responsive to an increase in concentration (n=8). In the case of heptyl benzene, an ANOVA having one factor within subjects, i.e., temperature (which, in fact, indicates headspace concentration) and one factor between subjects, i.e., unresponsive vs. responsive subjects regarding an increase in concentration, revealed significance for temperature (i.e., concentration) {$F(1,19)=10.364$, $p=0.0045$} and for the interaction temperature X group {$F(1,19)=34.358$, $p=0.0001$}. A comparable ANOVA for 2-tridecanone revealed significance for group {$F(1,19)=5.220$, $p=0.03$} and, although temperature slightly missed significance ($p=0.06$), the interaction temperature X group was significant {$F(1,19)=19.506$, $p=0.0003$}. For both chemicals, the results support the notion that one larger group of unresponsive subjects failed to increase their ocular detectability with an increase in concentration whereas another smaller group of responsive subjects did increase it.

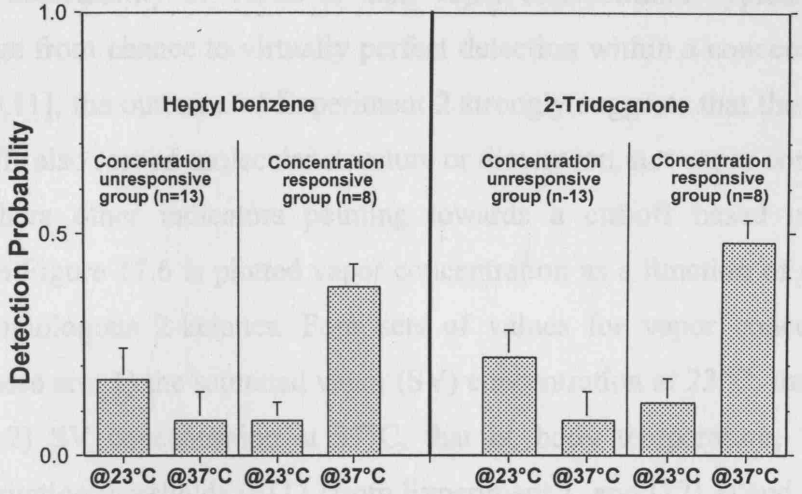


Figure 17.5. Comparison between a group of unresponsive (n=13) and a group of responsive (n=8) subjects in terms of their ability to precipitate or increase the ocular detectability of the VOC with an increase in concentration (achieved by increasing the temperature of the VOC). The left panel shows the results for heptyl benzene and the right panel those for 2-tridecanone. Bars indicate standard error. Data for the unresponsive groups represent the average of 260 judgments per stimulus; data for the responsive groups represent the average of 160 judgments per stimulus.

17.4. DISCUSSION

In the present work, a rational for the results obtained by Enrique Cometto-Muniz is now presented. The studies cited in the Introduction support the following notions: 1) chemesthesis is a receptor(s)-based process that, for most nonreactive VOCs, rests strongly on selective rather than specific interactions; 2) many and diverse receptors and neuron types are involved; 3) there is a cut-off effect in the chemesthetic ability of VOCs, and it seems to be related to molecular structure, size, and/or dimensions rather than to vapor concentration. The outcome of Experiment 1 identified heptyl benzene and 2-tridecanone as the cut-off homologs within their respective series. A previous investigation on detection of nasal pungency in anosmics, i.e., subjects lacking olfaction, found that they were unable to detect butyl benzene and higher alkylbenzenes, although the approach was not optimized to define a cut-off point [9]. The comparison of nasal vs. ocular cut-offs in chemesthesis from series of VOCs represents a fertile area of research to further unveil the basis for the phenomenon and warrants study on its own. The results of Experiment 2 showed that the cut-off could not be significantly overcome by an increase in vapor concentration of 5.0 times in the case of heptyl benzene or 6.9 times in that of 2-tridecanone. Considering that functions relating

chemesthetic detectability of VOCs to their vapor concentration typically cover the complete range from chance to virtually perfect detection within a concentration factor of just 10 [10,11], the outcome of Experiment 2 strongly suggests that the basis for the present cut-offs also rest on molecular structure or dimension, not vapor concentration.

Are there other indicators pointing towards a cut-off based on molecular dimension? In Figure 17.6 is plotted vapor concentration as a function of carbon chain length for homologous 2-ketones. Four sets of values for vapor concentration are presented. These are: 1) the saturated vapor (SV) concentration at 23°C, that is, at room temperature; 2) SV concentration at 37°C, that is, body temperature; 3) measured human eye irritation thresholds (EIT) {from Experiment 1, and [12] }; and 4) predicted, i.e., calculated, eye irritation thresholds from a QSAR model [13]. The figure shows that, although the model occasionally might underestimate (2-propanone) or overestimate (2-heptanone) a threshold, it is often right on target (2-pentanone, 2-nonanone, 2-dodecanone). More importantly, the equation relating threshold concentration to carbon chain length along ketones is virtually identical, in slope and in relative position, for measured and for calculated values (Figure 17.6). Both equations suggest a cut-off at the level of 2-tridecanone (or even 2-dodecanone) where the threshold clearly reaches the SV concentration at 23°C. Nevertheless, the SV concentration of 2-tridecanone at 37°C is still clearly above the threshold expected from either equation, indicating that if the cut-off at 23°C were concentration-based, it should be easily overcome by heating the VOC to 37°C. This did not happen, and so the conclusion is that the cut-off is not due to a physical mechanism but to a chemical mechanism.

An analogous comparison is presented in Figure 17.7 for homologous alkylbenzenes, where the sources for predicted and measured EIT are as cited above. Here, the values for measured and predicted EIT diverge progressively from toluene onwards, as shown by the lines representing the corresponding equations, in the direction of calculated values increasingly overestimating the measured threshold. As a result, the equation for calculated values predicts a cut-off at the level of propyl benzene for stimuli at 23°C and at the level of octyl benzene for stimuli at 37°C. In contrast, the equation for measured values suggests a cut-off at the level of heptyl benzene (or lower) for stimuli at 23°C but no cut-off, at least until well beyond octyl benzene, for stimuli at 37°C. The present analysis at 23°C agrees well with the trend in measured EIT, but fails to show that the cut-off is overridden when heptyl benzene is heated to 37°C as would

have been expected from a concentration-based limitation (that is, a physical mechanism). This suggests that the basis for the cut-off along alkylbenzenes is likely to be based on a chemical mechanism involving molecular dimensions (or size), as for the ketones.

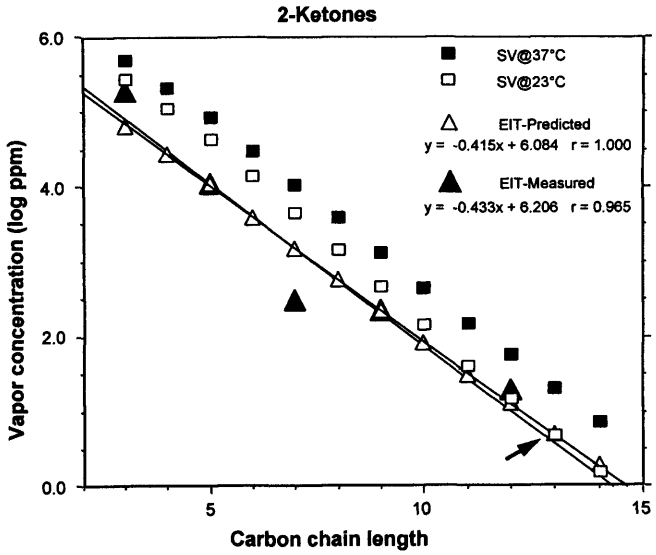


Figure 17.6. Relationship between various set of values for vapor concentration and carbon chain length for homologous 2-ketones. The set of values include: 1) saturated vapor (SV) concentration at 23°C; 2) SV concentration at 37°C {both sets from: extrapolated plots of log vapor pressure vs. carbon chain length and [14,15]}; 3) measured human eye irritation thresholds (EIT) {from Experiment 1, and [12]}; 4) predicted, i.e., calculated, eye irritation thresholds (EIT) {from [13]}.

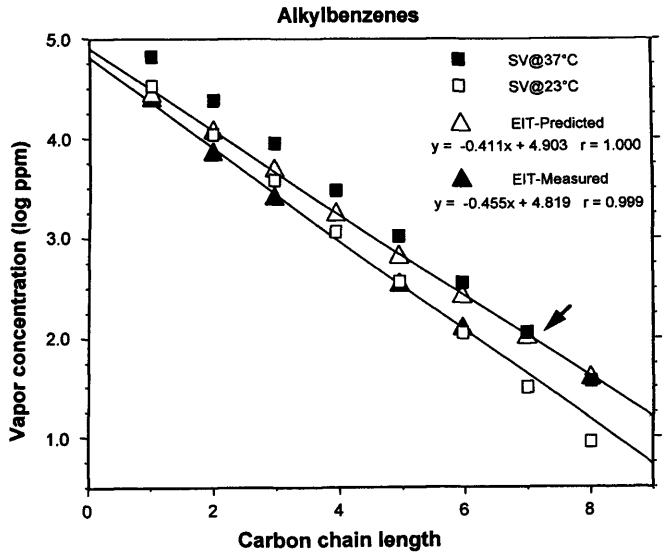


Figure 17.7. Analogous to Figure 6 but for homologous alkylbenzenes. Carbon chain length refers to the side chain attached to the benzene ring. Values of SV concentration at both temperatures from [16]; the rest as in Figure 6.

Despite the lack of a significant increase for the group data (n=21) in the detectability of 2-tridecanone and heptyl benzene at the higher concentration (Figure 17.3), the analysis of individual data (Figure 17.4) revealed that a sub-group of participants (n=8) did significantly increase their detection (Figure 17.5). A similar result was observed for the cut-off homologs decyl acetate and 1-undecanol in a study of ocular chemesthesis among acetate esters and n-alcohols [4]. Given that the various chemesthetic receptors putatively involved are all likely to be proteins [14], the outcome could reflect genetic variability [15,16] across subjects in the precise structure or dimension of receptor sites, leading to cut-offs displaced by one or two carbon chain units. These differences across subjects have also been observed in studies of nasal chemesthesis from homologous families of esters [17]. It can also be noted that differences in tear film break-up time represent another source of variability across individuals [18,19]. An additional source of individual variability could be the thickness of the tear film. Unfortunately, investigators have not reached a consensus on either the thickness of the film or its degree of variability across subjects [20,21, 22].

An analysis similar to that performed here led to the conclusion [4] that the cut-off previously reported for decyl acetate and 1-undecanol was unlikely to be based on a failure to reach a high enough vapor concentration. Instead, it was probably based on these molecules exceeding a critical size and becoming too large to be accommodated effectively by the necessary number and/or types of chemesthetic receptors. The present work incorporates heptyl benzene and 2-tridecanone, from homologous alkylbenzene and 2-ketone series, as additional cut-off stimuli, also likely to have surpassed some critical molecular dimension to evoke ocular chemesthesis.

In order to explore the possibility of there being some critical molecular dimension, beyond which ocular chemesthesis is not evoked, we have calculated the molecular dimensions of the compounds in the two homologous series. HyperChem software [23] was used to optimize the conformation of minimum energy of a VOC in the gas phase. The geometry optimization was carried out using the semi-empirical and accurate AM1 method [24], whereas the Polak-Ribiere method [25], a good general-purpose optimizer, was chosen as the minimization algorithm. Then the dimensions of the box into which the minimum energy conformation of the VOC will just fit can be obtained. The dimensions are X, Y and Z, corresponding to the width, depth and length of the box. We denote this box as the 'VOC box'. It is important to

note that the volume of the box, calculated as XYZ, is not the same as the volume of the particular VOC. This is why the box dimension is denoted just as XYZ.

This procedure was carried out for the VOCs in the homologous series of ketones and alkylbenzenes. There was nothing unusual with respect to any single one of the dimensions and a 'cut-off' point. The dimensions increase gradually along the homologous series with no divergence from regularity at or near heptyl benzene or 2-tridecanone. The above calculations refer to the gas phase, which can be taken as almost equivalent to a non-polar non-aqueous solvent. It is well established that in non-aqueous solvents proteins adopt an extended or non-folded conformation [26] whereas in water they adopt a folded conformation [27,28]. There is a lack of information on conformations of long chain homologues in solution, but many years ago Abraham [29] examined the solubility of a number of homologous series in water. He showed that the aqueous solubility decreased regularly with increase of chain length along any homologous series up to a certain point, after which the solubility decrease was much less. For the n-alkanes this point was reached near dodecane, and for the n-alcohols near to 1-dodecanol [29]. Whether or not the origin of these effects is folding of the alkyl chain in water, we think it unlikely to be coincidental that the cut-off point for ocular chemesthesis of the n-alcohols at 1-undecanol [4] is very close to the aqueous solubility effect at 1-dodecanol.

It therefore seemed important to repeat calculations on the homologous series, but this time for VOCs in an aqueous environment rather than in the gas phase. Calculations were therefore set up with a box containing 1264 water molecules, and the VOC placed in the center of the box. The size of this box, which determines the number of water molecules in it, was chosen based on the size of the largest VOC in the homologous series, which is 2-octadecanone, thus ensuring that even the largest VOC is surrounded by water molecules. The conformation of the VOC is then altered until it reaches the position of minimum energy. After the optimization the 'VOC box' was again constructed, as the smallest box enclosing the VOC solute. In this calculation, the atoms of the target VOC move during the structural optimization to their final minimum energy conformation. The surrounding water molecules do not move during the minimization, but produce a fixed potential field with which the VOC interacts. These calculations showed that for the 2-alkanones, the 'VOC box' XYZ value reaches a local minimum at 2-tetradecanone, and for the alkylbenzenes XYZ reaches another local minimum at octylbenzene.

The observation of local minima of XYZ at or near to the observed cut-off points in the homologous series was interesting enough to repeat the calculations, but this time allowing the water molecules to move during the optimization. Due to the large calculation time involved, we chose the molecular mechanics method MM+ [30], not as accurate as AM1, but much faster. The Polak-Ribiere method was again chosen as the minimization algorithm.

Results of the 'VOC box' dimensions, X, Y, and Z, for calculations in the gas phase using the AM1 procedure are in Tables 17.1 and 17.2, in units of Å and Å^3. The dimensions alter quite regularly with increase in the number of carbon atoms in the VOC.

Table 17.1
Dimensions (Å and Å^3) of the box enclosing the VOC for the 2-ketone series in vacuo

VOC	MF	X	Y	Z	XYZ
2-octanone	C8H16O	2.71916	1.81467	10.58340	52.22250
2-nonanone	C9H18O	2.74608	1.81441	11.79974	58.79238
2-decanone	C10H20O	2.70962	1.81501	13.07824	64.31862
2-undecanone	C11H22O	2.71031	1.81547	14.30010	70.36345
2-dodecanone	C12H24O	2.71518	1.81485	15.57532	76.74964
2-tridecanone	C13H26O	2.70229	1.81460	16.80535	82.40630
2-tetradecanone	C14H28O	2.72914	1.81491	18.07128	89.50964
2-pentadecanone	C15H30O	2.71644	1.81506	19.30550	95.18580
2-hexadecanone	C16H32O	2.75917	1.81501	20.55960	102.96086
2-heptadecanone	C17H34O	2.73566	1.81515	21.79995	108.25056
2-octadecanone	C18H36O	2.76912	1.81491	23.05823	115.88383

Table 17.2
Dimensions (Å and Å^3) of the box enclosing the VOC for the alkylbenzene series in vacuo

VOC	MF	X	Y	Z	XYZ
ethylbenzene	C8H10	4.61663	2.88829	7.28554	97.14660
propylbenzene	C9H12	4.63389	2.71558	8.38698	105.53923
butylbenzene	C10H14	4.76730	3.13234	9.48998	141.71202
pentylbenzene	C11H16	4.76817	2.76489	10.87958	143.43057
hexylbenzene	C12H18	4.90214	3.04592	11.90531	177.76445
heptylbenzene	C13H20	4.90771	3.45206	12.73413	215.73793

Table 17.2 *(continued)*

VOC	MF	X	Y	Z	XYZ
octylbenzene	C14H22	4.94832	2.56547	14.34670	182.12801
nonylbenzene	C15H24	4.87857	2.67489	15.75159	205.55255
decylbenzene	C16H26	5.46956	3.04489	16.26118	270.81708
undecylbenzene	C17H28	4.87246	3.81671	17.55405	326.44857
dodecylbenzene	C18H30	5.66281	3.50860	17.75894	352.84412

The corresponding results for the VOCs in aqueous solution using the MM+ procedure are in Tables 17.3 and 17.4. It is clear by inspection of Tables 17.3 and 17.4 that the VOC dimensions do not alter regularly with the number of carbon atoms in the VOC. Plots of Z and of XYZ against carbon chain length for the two homologous series reveal that the discontinuities are only evident in plots of XYZ, not Z, against carbon chain length, Figures 17.8 and 17.9. The plot of XYZ against carbon number for the 2-ketones, Figure 17.8, shows a discontinuity at C9 as well as at C15, and for the alkylbenzenes the plot of XYZ against carbon number, Figure 17.9, shows a discontinuity at the side chain carbon number C3 and C8, corresponding to a total carbon number of C9 (propylbenzene), and C14 (octylbenzene). The discontinuities in XYZ occur at positions along the homologous series very close to the observed cut-offs in ocular chemesthesis, that is at 2-tridecanone and heptylbenzene. These calculations refer to a completely non-biological system, but cut-off effects are indeed observed in such systems. The solubility phenomena observed by Abraham [29] is an indirect example. A direct example is the complexation of 1-alkanols with dipalmitoyl-L-α-phosphatidylcholine (DPPC) in D_2O/CCl_4 reversed micelles [31] that showed a very sharp cut-off at 1-tetradecanol. Examples of cut-offs in the interaction of aqueous homologous series with biological systems include anesthesia [2,32,33,34] and effects on enzymes [35,2,36].

The most interesting of the latter is the cut-off effects observed by Franks and Lieb [2] for the inhibition of activity of the firefly luciferase enzyme by aqueous alkanes and 1-alkanols. For the alkanes, a sharp but small cut-off point at hexane (C6) was observed, followed by another gradual cut-off around dodecane (C12). The 1-alkanols similarly showed a sharp small cut-off at 1-heptanol (C7) and a gradual but more severe cut-off around 1-dodecanol (C12). These twin cut-off effects are extraordinarily similar to those shown in Figures 17.8 and 17.9. Franks and Lieb suggested that the luciferase

binding site was amphiphilic, and contained polar as well as non-polar parts. The binding site can accommodate two molecules of an inhibitor provided that the latter is less than C6. For homologues higher than C6, only one molecule binds, but the site cannot accommodate even a single molecule higher than C12.

Table 17.3
Dimensions (Å and $Å^3$) of the box enclosing the VOC for the 2-ketone series in water

VOC	MF	X	Y	Z	XYZ
2-octanone	C8H16O	2.71848	1.99700	10.76183	58.42387
2-nonanone	C9H18O	2.91355	2.37967	12.01658	**83.31440**
2-decanone	C10H20O	2.72092	1.97314	13.31027	71.45959
2-undecanone	C11H22O	2.73122	2.15495	14.56509	85.72491
2-dodecanone	C12H24O	2.77172	1.96935	15.84842	86.50839
2-tridecanone	C13H26O	2.77476	2.03814	17.12012	96.82026
2-tetradecanone	C14H28O	2.74317	2.11536	18.39307	106.73116
2-pentadecanone	C15H30O	2.76085	2.66093	19.61415	**144.09395**
2-hexadecanone	C16H32O	2.70382	2.39481	20.89631	135.30643
2-heptadecanone	C17H34O	2.77047	2.25086	22.19663	138.41686
2-octadecanone	C18H36O	2.91829	2.29446	23.47993	157.21926

Table 17.4
Dimensions (Å and $Å^3$) of the box enclosing the VOC for the alkylbenzene series in water

VOC	MF	X	Y	Z	XYZ
Ethylbenzene	C8H10	4.77286	1.89689	7.34661	66.51320
Propylbenzene	C9H12	4.72010	2.63893	8.45265	**105.28632**
butylbenzene	C10H14	4.86036	1.87069	9.80521	89.15119
pentylbenzene	C11H16	4.85192	1.82110	11.02503	97.41531
hexylbenzene	C12H18	4.90568	1.97422	12.28468	118.97579
heptylbenzene	C13H20	4.90900	1.91591	13.51293	127.09184
octylbenzene	C14H22	4.85982	2.74466	14.76311	**196.91853**
nonylbenzene	C15H24	4.92206	2.11115	16.04979	166.77669
decylbenzene	C16H26	4.90253	2.45356	17.25864	207.59817
undecylbenzene	C17H28	4.92550	2.30808	18.46164	209.88020
dodecylbenzene	C18H30	4.82230	2.84850	19.74669	271.24688

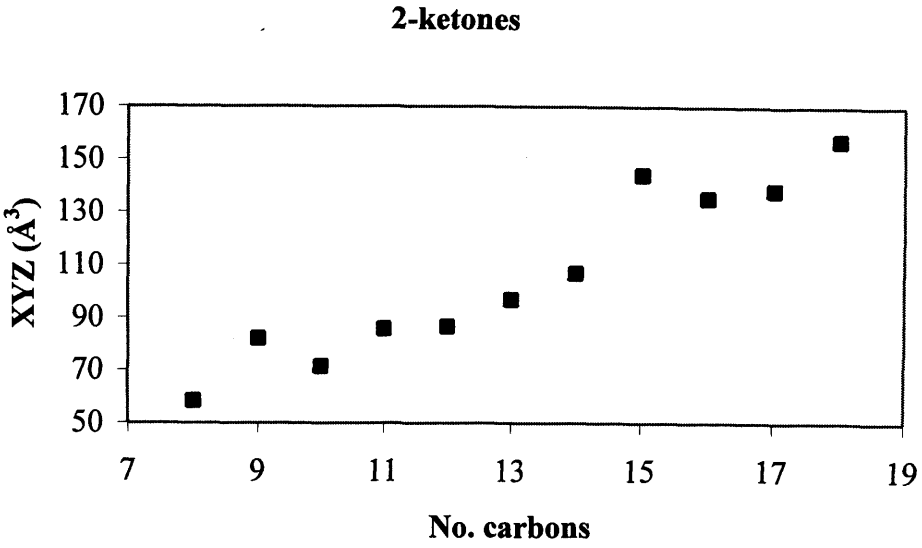


Figure 17.8. A plot of the VOC box dimension, XYZ, against the number of carbon atoms for the 2-alkanone series in water.

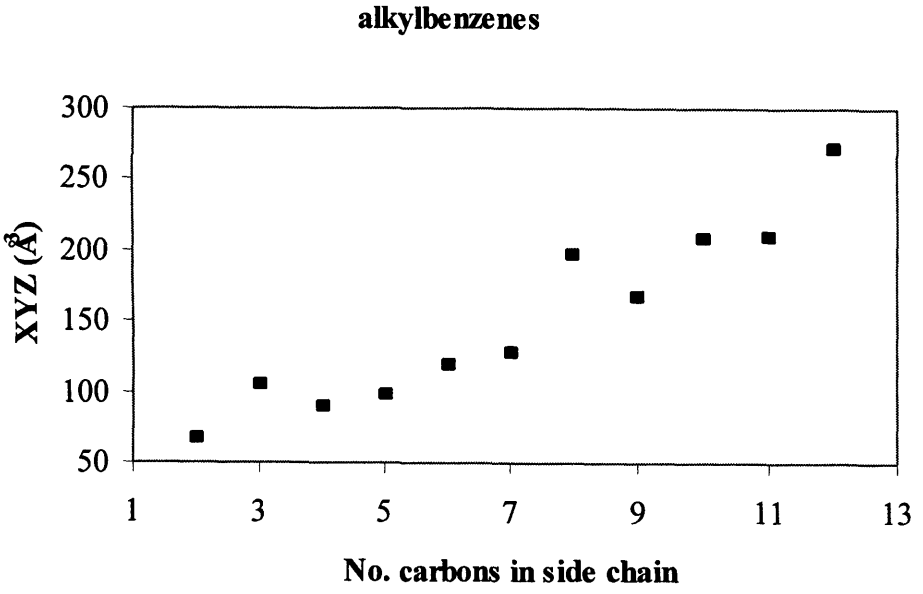


Figure 17.9. A plot of the VOC box dimension, XYZ, against the number of carbon atoms in the alkyl side chain for the alkylbenzene series in water.

Whether these considerations apply to ocular chemesthesis is another matter. However the present calculations strongly supports the conclusion that the observed lack of ocular chemesthesis in the higher homologues is due to a chemical effect that is connected to the size of the VOC and the receptor, and not to the low vapor pressures of the higher homologues. Furthermore, if the present results of calculations are found to

be general, then they constitute a method for the prediction of the cut-off point in any homologous series. This will be explored by extending calculations to other homologous series.

The mathematical model constructed for the correlation of eye irritation thresholds has been viewed [37, 13, 38, 4] as though the receptor was in a largely non-aqueous (but quite polar) environment, which seems, at first sight, to be contrary to the present results. Franks and Lieb [33], however, have argued that anesthetics act on hydrophobic protein pockets that are exposed to water. This would be the case in eye irritation, because the pre-ocular film that lies adjacent to the epithelium, the tear film, is largely composed of water [39]. Hence the receptor could act as somewhat polar non-aqueous medium, capable of interacting with hydrophobic parts of VOCs in a homologous series, and yet be responsive to configurational changes in the VOC brought about by the adjacent aqueous medium.

17.5. REFERENCES

1. J. E. Cometto-Muñiz, (2001) Physicochemical basis for odor and irritation potency of VOCs. In *Indoor Air Quality Handbook* (J. D. Spengler, J. Samet and J. F. McCarthy, Eds.), pp. 20.21-20.21. McGraw-Hill, New York.
2. N. P. Franks, W. R. Lieb, Nature 316 (1985) 349.
3. J. E. Cometto-Muñiz, W. S. Cain, M. H. Abraham, Chem. Senses 30 (2005) 627.
4. J. E. Cometto-Muñiz, W. S. Cain, M. H. Abraham, Toxicol. Appl. Pharmacol. 207 (2005) 232.
5. W. S. Cain, Ear Nose Throat J. 68 (1989) 316.
6. J. E. Cometto-Muñiz, W. S. Cain, M. H. Abraham, Chem. Senses 28 (2003) 467.
7. D. Podlekareva, Z. Pan, S. Kjaergaard, L. Molhave, Int. Arch. Occup. Environ. Health 75 (2002) 359.
8. N. A. Macmillan, C. D. Creelman, (1991) *Detection theory: A user's guide.* Cambridge University Press, Cambridge.
9. J. E. Cometto-Muñiz, W. S. Cain, Am. Ind. Hyg. Assoc. J. 55 (1994) 811.
10. W. S. Cain, R. A. de Wijk, A. A. Jalowayski, G. Pilla Caminha, R. Schmidt, Indoor Air 15 (2005) 445.
11. J. E. Cometto-Muñiz, W. S. Cain, M. H. Abraham, J. M. R. Gola, J. Appl. Toxicol. 22 (2002) 25.

12. J. E. Cometto-Muñiz, W. S. Cain, Chem. Senses 20 (1995) 191.
13. M. H. Abraham, J. M. R. Gola, J. E. Cometto-Muñiz, W. S. Cain, Indoor Built Environ. 10 (2001) 252.
14. D. Julius, A. I. Basbaum, Nature 413 (2001) 203.
15. G. I. Elmer, J. O. Pieper, S. S. Negus, J. H. Woods, Pain 75, (1998) 129.
16. J. S. Mogil, S. G. Wilson, K. Bon, S. E. Lee, K. Chung, P. Raber, J. O. Pieper, H. S. Hain, J. K. Belknap, L. Hubert, G. I. Elmer, J. M. Chung, M. Devor, Pain 80 (1999) 67.
17. W. S. Cain, N.-S. Lee, P. M. Wise, R. Schmidt, B.-H. Ahn, J. E. Cometto-Muñiz, M. H. Abraham, Physiol. Behav. 88 (2006) 317
18. J. J. Nichols, K. K. Nichols, B. Puent, M. Saracino, G. L. Mitchell, Optom. Vis. Sci. 79 (2002) 363.
19. R. Stodtmeister, T. Christ, W. Gaus, *Klin. Monatsbl. Augenheilkd.* 183 (1983) 485.
20. P. E. King-Smith, B. A. Fink, R. M. Hill, K. W. Koelling, J. M. Tiffany, Curr. Eye Res. 29 (2004) 357.
21. J. I. Prydal, P. Artal, H. Woon, F. W. Campbell, Invest. Ophthalmol. Vis. Sci. 33 (1992) 2006.
22. C. J. Radke, Curr. Eye Res. 30 (2005) 1131-1132; author reply 1133-1134.
23. HyperChem. Version 7.5. Hypercube Inc., 1115 NW 4th Street, Gainesville, Florida 32601, USA.
24. M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, J. J. P. Stewart, J. Am. Chem. Soc. 107 (1985) 3902.
25. A. Schwartz, E. Polak, J. Optimiz. Theory App. 92 (1997) 1.
26. C. M. Soares, V. H. Teixeira, A. M. Baptista,. Biophys. J. 84 (2003) 1628.
27. R. Glockshuber, Chimia 55 (2001) 281.
28. J. M. Yon, Cell. Mol. Life Sci. 53 (1997) 557
29. M. H. Abraham, J, Chem. Soc., Faraday Trans. 80 (1984) 153
30. HyperChem Computational Chemistry (2002). pp. 190-197. Hypercube Inc., USA.
31. J. S. Chiou, S. M. Ma, H. Kamaya, I. Ueda, Science 248 (1990) 583.
32. N. P. Franks, W. R. Lieb, Environ. Health Perspect. 87 (1990) 199.
33. N. P. Franks, W. R. Lieb, Nature 367 (1994) 607.
34. Y. Katz, I. Aharon, Bull. Math. Biol. 62 (2000) 1.
35. A. N. Fonteh, K. McBride, W. A. Gibbons, Biochem. Soc. Trans. 17 (1989) 721.
36. D. McKenzie, N. P. Franks, W. R. Lieb, Br. J. Pharmacol. 115 (1995) 275

37. M. H. Abraham, R. Kumarsingh, J. E. Cometto-Muñiz, W. S. Cain, Toxicol. in Vitro 12 (1998) 403.
38. M. H. Abraham, M. Hassanisadi, M. Jalali-Heravi, T. Ghafourian, W. S. Cain, J. E. Cometto-Muñiz, Toxicol. Sci. 76 (2003) 384.
39. S. Iwata, Int. Ophthalmol. Clin. 13 (1973) 29.

Chapter 18 Suggestions for Future Work

Future work should clearly involve the determination of Abraham descriptors for more indoor air pollutants (IAPs) with the aim of improving the Abraham Solvation Equation for Odour Detection Threshold (ODT) and Nasal Pungency Threshold (NPT). Of course, it is also necessary to increase the number of measurements of ODT and NPT for more IAPs to obtain these two improved equations.

However, a more important piece of work would be to obtain a full equation for the psychometric constants (C and D), which as mentioned in a previous chapter provide much more information than what a single "threshold value" can convey, for instance on how the senses of smell and chemesthesis process mixtures of compounds. The number of chemicals tested in this way at the moment is too low, and consequently the equations obtained for C and D in the two chapters on psychometric functions in this works can be improved. It must surely be the case that the number of descriptors present in the equations for C and D will increase as the number of compounds studied with different properties increases.

On the other hand, the study of cut off points on homologous series should continue in order to reach a firm conclusion on whether these cut off points really depend on molecular dimension (structure-based) as suggested by all the studies carried out until now, or whether they depend on vapour concentration (concentration-based). It is necessary to demonstrate that the results shown in this work are not just by chance but are real. To do so, it would be very convenient to carry out molecular dimension analysis with other computer programs, in the same way as was done with Hyperchem in this work, to compare results, and to see whether the results are similar or not. And finally, see if an adequate conclusion can be reach with these analyses. Furthermore, this study should be extended to cover the areas of nasal pungency and odour detection, and not just the cut off phenomenon in eye irritation.

The last proposal is to carry out measurements on sensory irritation and odour, using environmental chambers. These chambers allow people to be in contact with the chemicals in a more similar way as happens in a typical day spent in a house, or a working place. They allow people to be in contact with the stimulus not just for a few seconds, but for a longer, more natural, time. The main problems with the use of

environmental chambers is the increase in expense, as well as the length of time that the experiments take.

Another important aspect in the future would be to investigate the possibility of devising a software package which allows the automatic calculation of the Abraham descriptors from experimental data. How? The first thing is to provide a package with a very simple interface that allows anyone to use it. The package would have to contain a database with all the Abraham equations that are already known for water-solvent and gas-solvent systems. It should give the possibility of introducing the equations experimentally obtained in the laboratory and of making a larger database to be used in each particular case. Obviously, the software should give the option of using this modified database to calculate the descriptor instead of the one that comes by default.

This application should work in the same way as Solver, which minimises the sum of squares on the required equations to fit the targeted cells, which contain the descriptors to be calculated (usually **S**, **A**, **B** and **L**). The values are accepted when the overall sums of squares are at a minimum. The proposed application should come with the option to calculate the **E** descriptor just introducing the refractive index, or if known, introducing its value manually. This descriptor is then fixed, and only the other four descriptors are calculate by the software. Otherwise, this **E** descriptor should be introduced as one of the parameter to be calculated by the program. Therefore, the application should offer the chance to choose the number of descriptors to be calculated and indicate which of the five descriptors are allowed to float, and which are fixed. Such an application would reduce the time and effort spent in continually calculating, and sometimes recalculating, the descriptors.